BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N

A2

(11) International Publication Number: **WO 97/18454**

(43) International Publication Date: 22 May 1997 (22.05.97)

(21) International Application Number: PCT/US96/18567

(22) International Filing Date: 15 November 1996 (15.11.96)

(30) Priority Data:

60/006,838	16 November 1995 (16.11.95)	US
08/594,031	30 January 1996 (30.01.96)	US

(71)(72) Applicant and Inventor: THOMPSON, Timothy [US/US]; 4835 Jason Street, Houston, TX 77096 (US).

(74) Agents: REMENICK, James et al.; Baker & Botts, L.L.P., The Warner, 1299 Pennsylvania Avenue, N.W., Washington, DC 20004 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: METHOD FOR IDENTIFYING METASTATIC SEQUENCES

(57) Abstract

The invention relates to methods for the identification of metastatic sequences. Cells from a cell line or an animal tissue are treated to form a cell line predisposed to metastasis. Treated cells are implanted in an animal of a primary site and incubated for a period of time sufficient for the cells to proliferate and develop metastases at secondary sites. Expressed sequences from cells at the primary and secondary sites are amplified by differential display polymerase chain reaction and compared. Differentially expressed sequences are identical and can be cloned and sequenced. These sequences can be used as probes in the diagnosis of metastatic disorders, as probes to isolate metastatic sequences and as a therapeutic agent.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

METHOD FOR IDENTIFYING METASTATIC SEQUENCES

Rights in the Invention

This invention was made in part with United States Government support under grant number CA350129, awarded by the National Cancer Institute, National Institute of Health and the United States Government has certain rights in the invention.

Background

1. Field of the Invention

The present invention relates to methods for the identification and isolation of metastatic sequences, to diagnostic probes and kits which contain metastatic sequences and to therapeutic treatments for neoplastic disorders based on metastatic sequences.

2. Description of the Background

The development of higher organisms is characterized by an exquisite pattern of temporal and spatially regulated cell division. Disruptions in the normal physiology of cell division are almost invariably detrimental. One such type of disruption is cancer, a disease that can arise from a series of genetic events.

Cancer cells are defined by two heritable properties, uncontrolled growth and uncontrolled invasion of normal tissue. A cancerous cell can divide in defiance of the normal growth constraints in a cell leading to a localized growth or tumor. In addition, some cancer cells also gain the ability to migrate away from their initial site and invade other healthy tissues in a patient. It is the combination of these two features that make a cancer cell especially dangerous.

An isolated abnormal cell population that grows uncontrollably will give rise to a tumor or neoplasm. As long as the neoplasm remains in a single location, it is said to be benign, and a complete cure may be expected by removing the mass surgically. A tumor or neoplasm is counted as a cancer if it is malignant, that is, if its cells have the ability to invade surrounding tissue. True malignancy begins when the cells cross the basal

lamina and begin to invade the underlying connective tissue. Malignancy occurs when the cells gain the ability to detach from the main tumor mass, enter the bloodstream or lymphatic vessels, and form secondary tumors or metastases at other sites in the body. The more widely a tumor metastasis, the harder it is to eradicate and treat.

As determined from the epidemiological and clinical studies, most cancers develop in slow stages from mildly benign into malignant neoplasms. Malignant cancer usually begins as a benign localized cell population with abnormal growth characteristic called a dysplasia. The abnormal cells acquire abnormal growth characteristics resulting in a neoplasia characterized as a cell population of localized growth and swelling. If untreated, the neoplasia *in situ* may progress into a malignant neoplasia. Several years, or tens of years may elapse from the first sign of dysplasia to the onset of full blown malignant cancer. This characteristic process is observed in a number of cancers. Prostate cancer provides one of the more clear examples of the progression of normal tissue to benign neoplasm to malignant neoplasm.

The walnut-sized prostate is an encapsulated organ of the mammalian male urogenital system. Located at the base of the bladder, the prostate is partitioned into zones referred to as the central, peripheral and transitional zones, all of which surround the urethra. Histologically, the prostate is a highly microvascularized gland comprising fairly large glandular spaces lined with epithelium which, along with the seminal vesicles, supply the majority of fluid to the male ejaculate. As an endocrine-dependent organ, the prostate responds to both the major male hormone, testosterone, and the major female hormones, estrogen and progesterone. Testicular androgen is considered important for prostate growth and development because, in both humans and other animals, castration leads to prostate atrophy and, in most cases, an absence of any incidence of prostatic carcinoma.

The major neoplastic disorders of the prostate are benign enlargement of the prostate, also called benign prostatic hyperplasia (BPH), and prostatic carcinoma; a type of neoplasia. BPH is very common in men over the age of 50. It is characterized by the presence of a number of large distinct nodules in the periurethral area of the prostate. Although benign and not malignant, these nodules can produce obstruction of the urethra causing nocturia, hesitancy to void, and difficulty in starting and stopping a urine stream upon voiding the bladder. Left untreated, a percentage of these prostate hyperplasia and neoplasias may develop into malignant prostate carcinoma.

In its more aggressive form, transformed prostatic tissues escape from the prostate capsule and metastasize invading locally and throughout the bloodstream and lymphatic system. Metastasis, defined as tumor implants which are discontinuous with the primary tumor, can occur through direct seeding, lymphatic spread and hematogenous spread. All three routes have been found to occur with prostatic carcinoma. Local invasions typically involve the seminal vesicles, the base of the urinary bladder, and the urethra. Direct seeding occurs when a malignant neoplasm penetrates a natural open field such as the peritoneal, pleural or pericardial cavities. Cells seed along the surfaces of various organs and tissues within the cavity or can simply fill the cavity spaces. Hematogenous spread is typical of sarcomas and carcinomas. Hematogenous spread of prostatic carcinoma occurs primarily to the bones, but can include massive visceral invasion as well. It has been estimated that about 60% of newly diagnosed prostate cancer patients will have metastases at the time of initial diagnosis.

Surgery or radiotherapy is the treatment of choice for early prostatic neoplasia. Surgery involves complete removal of the entire prostate (radical prostatectomy), and often removal of the surrounding lymph nodes, lymphadenectomy. Radiotherapy, occasionally used as adjuvant therapy, may be either external or interstitial using ^{125}I. Endocrine therapy is the

treatment of choice for more advanced forms. The aim of this therapy is to deprive the prostate cells, and presumably the transformed prostate cells as well, of testosterone. This is accomplished by orchiectomy (castration) or administration of estrogens or synthetic hormones which are agonists of luteinizing hormone-releasing hormone. These cellular messengers directly inhibit testicular and organ synthesis and suppress luteinizing hormone secretion which in turn leads to reduced testosterone secretion by the testes. Despite the advances made in achieving a pharmacologic orchiectomy, the survival rates for those with late stage carcinomas are rather bleak.

Summary of the Invention

The present invention overcomes the problems and disadvantages associated with current strategies and designs and provides new methods for the identification of sequences related to metastasis.

One embodiment of the invention is directed to methods for the identification of a metastatic sequence. One or more oncogenic sequences are transfected into a cell to form a transfected cell. The transfected cell is introduced into a primary site of a host animal to establish a colony which is incubated in the animal for a period of time sufficient to develop both a primary tumor and a malignant tumor. Expressed sequences are harvested from the primary tumor and the metastasis. Harvested sequences are compared to each other and to non-metastatic cells to identify sequences related to metastasis. Dominant metastatic genes are genes whose expression leads to metastasis. Such genes are typically expressed at high levels in metastatic cells and not significantly expressed in normal or nonmetastatic cells. Recessive metastatic genes, genes whose expression prevents metastasis, may be selectively expressed in normal and nonmetastatic cells and absent in metastatic cells. Dominant and recessive metastatic genes may act directly or act pleiotropically by enhancing or

inhibiting the expression or function of other dominant and recessive metastatic genes.

Another embodiment of the invention is directed to methods for identifying metastatic sequences. A mammalia cell is treated with a metastatic agent and the treated cell is implanted into a primary site of a host mammal. The host animal is maintained for a period of time sufficient for the cells to proliferate and to develop a metastatsis at a secondary site. Expressed squences from cells of the primary cite and cells of the secondary site are reverse transcribed into cDNA by differential display polymerase chain reaction to identify differentially expressed sequences.

Another embodiment of the invention is directed to sequences isolated by the methods of the invention. Sequences may be in the form of DNA, RNA or PNA. The nucleic acid may be single-stranded or double-stranded. Single stranded nucleic acid may be in the form of a sense strand or an antisense strand. In addition, the sequence may be part of a homologous recombination vector designed to recombine with another metastatic sequence.

Another embodiment of the invention is directed to a method for treating a neoplastic disorder comprising administering a pharmaceutically effective amount of a metastatic nucleic acid to a patient. The nucleic acid may be single-stranded in the sense or the antisense direction. Alternatively, the nucleic acid may be packaged in a viral vector such as, for example, a retroviral, a vaccinia or an adenoviral vector. Administration may be performed by injection, pulmonary absorption, topical application or delayed release of the nucleic acid along with a pharmaceutically acceptable carrier such as water, alcohols, salts, oils, fatty acids, saccharides, polysaccharides and combinations thereof.

Another embodiment of the invention is directed to a kit for detecting of the presence or absence of a metastatic sequence.

Other objects and advantages of the invention are set forth in part in the description which follows, and in part, will be obvious from this description, or may be learned from the practice of the invention.

Description of the Drawings

Figure 1 Schematic showing two paths in the multistep progression to cancer.

Figure 2 Staining of primary tumor (A) and metastatic deposit (B) from the lung of the same animal

Figure 3 Staining of normal human prostate (A), moderately differentiated human prostate tumor (B and C), and poorly differentiated prostate tumor (D).

Figure 4 Schematic of method for isolating a metastatic gene from a gene ablated mouse strain.

Figure 5 Schematic showing method to establish a tumor and a metastatic transplant from fetal tissue(A) and from cell lines and tumors (b).

Figure 6 Isolation and characterization of nmb gene expression by DD-PCR and RNA blot in primary and metastatic cells.

Figure 7 Differential expression of multiple genes is determined by DD-PCR and RNA blot of primary and metastatic cells.

Figure 8 Caveolin identified as a differentially expressed gene by DD-PCR.

Figure 9 Differential expression of genes isolated by DD-PCR confirmed by RNA blots.

Figure 10 RNA blot analysis of total tumor mRNA using clone 29 GADPH probes.

Figure 11 RNA blot of three independent MPR metastatic tumors and 5 MPR non-metastatic tumors.

Figure 12 Nucleotide sequences of metastatic nucleic acids.

Figure 13 Characterization of metastatic sequences isolated.

Figure 14 Immunohistological staining of primary and metastatic human prostate tumors using anti-caveolin antibodies.

Description of the Invention

As embodied and broadly described herein, the present invention is directed to methods for identifying metastatic sequences, to the metastatic sequences identified, to methods for the detection, diagnosis and treatment of disorders related to metastasis, and to diagnostic kits which comprise these sequences.

The ability of cancers to metastasize makes tumors difficult to eradicate by any means. Malignant cancer involves a multistage progression from, for example, normal tissue through hyperplasia, early adenoma, early carcinoma and finally to a metastatic tumor (Figure 1). Cells of a typical tumor loosen their adhesion to their original cellular neighbors and cross the basal lamina and endothelial lining to enter the body's circulation. Once in circulation, the metastatic cell exits from the circulation to disseminate throughout body and proliferate in a new environment.

Like the initial oncogenic event, the ability of a cell to metastasize requires additional mutationic or epigenetic changes. An understanding of the molecular mechanisms of metastasis allow for the design of treatments to inhibit metastasis. Knowledge of stage specific gene expression for neoplastic disorders allows for early detection and typing of tumors. With early detection and typing, proper treatment may be administered to a patient with the neoplastic disorder earlier, which will lead to a higher probability of a complete cure.

For human prostate tumors, the study of stage specific tumors is difficult, if not impossible, as cell lines are extremely difficult to grow and it is rare that tissue becomes available from the primary tumor as well as metastatic disease from the same patient. This problem is exacerbated

because of the infrequent biopsy of metastatic deposits in concordance of isolation of material from the primary tumor. Furthermore, the growth of cell lines from malignant prostates has proved to be problematic over the last few decades. This is evidenced by the lack of cell lines from prostate cancer obtained under any conditions.

One embodiment of the invention is directed to a method for identifying a metastatic sequence. A mammalian cell is transformed into a pre-neoplastic or neoplastic state or phenotype by transfection with one or more oncogenic sequences. Alternatively, or in addition to transfection, the mammalian cell may be treated with an agent or subjected to a condition that potentiates the metastatic character of the cell or predisposes the cell to metastasis. The transfected or treated cell is implanted into a host animal at a primary site and grown for a period of time sufficient to develop a metastasis at a secondary site. Expressed sequences from cells of the primary site and cells at the secondary site are amplified by differential display polymerase chain reactions. PCR products from these reactions are compared and the metastatic sequence identified by alteration in the levels or patterns of the resulting products.

Mammalian cells from a wide variety of tissue types and species are suitable for transfection or treatment including surgically obtained or primary or immortalized cells and cell lines. Cells may be from humans or primates, mice, rats, sheep, cows, rabbits, horses, pigs or guinea pigs or from transgenic or xenogeneic host mammals. Cells may be obtained from adult, juvenile or fetal tissue, and used directly from the mammal, from cryogenically preserved samples, or after culturing *in vitro* or *in vivo* for a period of time. *In vitro* culturing typically involves tissue culture conditions (*e.g.* 37°C; 5% CO_2) while *in vivo* culturing may involve successive passage of cells through host animals such as, for example, mice or rabbits. Cells passed *in vivo* may be obtained from sites proximal or distal to the site of implantation. The tissue type from which the cells are derived or obtained

may be any tissue which is susceptible to transfection or other treatment including, for example, urogenital tissues, epithelial cells, hepatic cells, fibroblasts lymphatic tissues, hematopoietic cells, cells of the immune system, cells of the gastrointestinal system and cells of the nervous system.

Cell types useful for the identification of metastatic sequences related to prostrate cancer include cells and cell lines of the fetal prostate lineage from normal or transgenic animals, and cells from normal or reconstituted prostate tissue. One method of generating reconstituted prostate cells is to isolate fetal prostate tissue and microdissect the fetal prostate epithelium away from fetal mesenchyme. Fetal prostate epitheliums may be genetically manipulated before reassociation with fetal mesenchyme (Figure 5A). Genetic manipulation involves treatment or transfection with a metastatic agent or a nucleic acid sequence that affects neoplastic or metastatic potential of the cell. Reassociation of fetal epithelium and mesenchyme is performed by implanting epithelium tissue within a pocket of mesenchyme tissue. After manipulation, cells are reimplanted into a mammalian host in a similar manner as other cells, such as reimplantation into or under the renal capsule.

Mammalian cells may be transfected by a variety of techniques, all of which are well-known to those of ordinary skill. Direct methods involve the introduction of genetic material into the nucleus of a cell by injection. These techniques include high velocity projectile injection, microinjection, and electroporation. Indirect methods, involving the active or passive uptake of the genetic information by the cell. Indirect techniques include transduction with recombinant vectors, and chemical or physical treatments such as calcium phosphate uptake, lipofection or dextran sulfate transfection. Chemical techniques rely on chemical carriers to introduce nucleic acids into a cell. These methods, for example, utilize unilamellar phospholipid vesicles (*e.g.* liposomes) loaded with DNA (or RNA). The approach relies on the fusion of the DNA containing vesicles with the

plasma membrane of the recipient cells. After entry, DNA traverse the cytoplasm and enter the nucleus. Another lipofection technique uses a synthetic cationic lipid such as N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTMA). DOTMA spontaneously associates with nucleic acids and forms unilamellar vesicles upon sonication. Genetic material is incorporated into these vesicles and subsequently transfected into the cell. Calcium phosphate co-precipitation involves mixing of purified nucleic acid with buffers containing phosphate and calcium chloride which results in the formation of a fine precipitate. Presentation of this precipitate to cells results in incorporation of the nucleic acid into cellular genome. Other chemicals, such as DEAE dextran or polybrene, when present in media with nucleic acids, can also cause the transfection of mammalian cells.

Physical methods of transfection rely on electric fields, needles and particles to enable nucleic acids to traverse the cellular membrane. Electric field mediated DNA transfection, commonly called electroporation, is based on the principle that membranes, when subjected to an electric field, undergo a reversible breakdown resulting in pores large enough to permit the passage of nucleic acids. In micro-projectile mediated gene transfer, micro-projectiles of subcellular dimensions are coated with nucleic acid and propelled at high velocity into a cell using a particle gun. The nucleic acid is introduced into the nucleus directly when the particles impinge upon the nucleus. In microinjection, nucleic acid is injected directly into the nucleus of a cell with a needle. Lasers have also been used to introduce minute holes in cellular membrane to allow introduction of nucleic acids. All these methods may be used for transfection and the selection of the method will depend on the cell type, the desired transfection efficiency and the equipment available.

The efficiency of transfection may be monitored and enhanced by the co-transfection of a selectable marker. If a marker is co-transfected

with a genetic construct, positively transformed cells may be separated from nontransformed cells by chemical selection. The efficiency of transfection will be increased in most cases because the chemicals will selectively kill non-transfected cells. The number of transfected cells may also be monitored by analyzing the degree of chemical resistance of the transfected cells. Markers commonly used for selection purposes include, for example, nucleic acids encoding dihydrofolate reductase, metallothionein, CAD, adenosine deaminase, adenylate deaminase, UMP synthetase, IMP 5'-dehydrogenase, xanthine-guanine phosphoribosyltransferase, mutant thymidine kinase, mutant HGPRTase, thymidylate synthetase, P-glycoprotein 170, ribonucleotide reductase, glutamine synthetase, asparagine synthetase, arginosuccinate synthetase, ornithine decarboxylase, HMG-CoA reductase, N-acetylglucosaminyl transferase, theronyl-tRNA synthetase, sodium or potassium dependent ATPase or derivatives or mutants of these nucleic acids. Markers may be used individually or in combination. Chemicals useful for selection include methotrexate, cadmium, PALA, Xyl-A, adenosine, 2'-deoxycoformycin, adenine, azaserine, coformycin, 6-azauridine, pyrazofuran, mycophenolic acid, limiting xanthine, hypoxanthine, aminopterin, thymidine, 5-fluorodeoxyuridine, adriamycin, vincristine, colchicine, actinomycin D, puromycin, cytocholasin B, emetine, maytansine, Bakers' antifolate, aphidicolin, methionine sulfoximine, β-aspartyl hydroxamate, albizziin, canavanine, α-difluoromethylornithine, compactin, tunicamycin, borrelidin, ouabain, and derivatives and analogs and combinations of these chemicals. Some chemicals, such as methotrexate, may be used individually while other chemicals, such as HAT (hypoxanthine, aminopterin and thymidine), need to be used in combination to be effective.

The oncogene transfection efficiency, the fraction of live cells tranfected by an oncogene, may be indirectly enhanced by chemical selection for a co-transfected marker. An oncogene is a sequence which can

predispose, or induce the cell into a pre-neoplastic or neoplastic condition or otherwise enhance the metastatic potential of the cell. Sequences with these properties are referred to as oncogenes and include *abl, ahi, akt, bcl, crk, dsi, erb, ets, evi, fes/fps, fim, fis, fgr, flv, fms, fos, gin, gli, int, jun, kit, mas, lck, met, mil/raf, mis, mlv, mos, myb, myc, neu, onc, pim, raf, ras, rel, ros, seq, sis, ski, spi, src, tcl, thy, trk,* and *yes.* Some oncogenes, such as *ras,* are oncogenic when mutated. Other oncogenes, such as *myc,* are oncogenic when overexpressed or underexpressed. Many oncogenes represent members of multigene families or homologs families. Homologs are proteins that have similar primary, secondary or tertiary structures. Genes may differ in nucleic acid sequence or encoded peptide sequence and still be homologs when the encoded polypeptides have similar spatial folding. Many oncogenes can be classified into dominant oncogenes and recessive oncogenes. One or more dominant oncogenes can confer a neoplastic or pre-neoplastic phenotype to a cell. One or more recessive oncogenes, when silenced, may also confer a neoplastic or preneoplastic phenotype. Gene silencing is performed by transfecting cells with nucleic acids which cause genetic ablation or by antisense suppression.

While any oncogene may be used, the preferred oncogenes are those that are normally associated with metastasis such as a metastasis specific gene. Such genes include for example, *TGF-β1, Cyclin D1 p21, p34, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb or α-actinin 3.* Metastatic-specific genes may be used individually or in combination with other oncogenes.

The metastatic potential of a cell may be altered, for example, by gene ablation with a sequence specific for a recessive oncogene. Recessive oncogene are those genes which encode products which can suppress oncogenesis and metastasis. A gene ablation sequence can be designed to specifically suppress a recessive oncogene. Ablation may include pre-transcriptional inhibition such as homologous recombination

with endogenous recessive oncogenes and post transcriptional inhibition such as the expression of antisense oncogenes to suppress translation. Gene ablation sequences may be targeted towards well known recessive oncogenes such as, for example, the retinoblastoma gene (Rb) or Bcg. Other candidates for ablation include metastatic genes previously isolated by the invention such as, for example, TGF-β1, cyclin D1, p21, p34, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb or α-actinin-3. The effects of ablating a recessive oncogene may include oncogenesis and metastases.

Alternatively, or in addition to transfecting the mammalian cell may be treated with an agent, either before or after transfection, that alters the expression of the cell's nucleic acids. Treatment may comprise contacting the cells with one or more agents which affect the neoplastic (*e.g.* neoplastic agents; phorbol esters), metabolization (*e.g.* metabolic agents), metastatic (*e.g.* metastatic agents), differentiation (*e.g.* differentiation agents; retinoic acid), activation or proliferation (*e.g.* growth factors) of the cell. Agents which can alter gene expression include chemicals such as benzanthracene (BA), dimethyl benzanthracene (DMBA) or 5-azacytidine. Alternatively, treatment may also comprise altered conditions such as hypoxia which involves subjecting a cell to a reduced oxygen content, exposable to radiation or other stresses to the cell.

Treatment may be *in vitro* or *in vivo* and may include for example, direct or indirect induction or suppression of well know oncogenic sequences and genes isolated by the invention such as, for example, *TGF-β1, Cyclin D1, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb, α actinin 3, and p34.* Gene expression induction includes transfecting expression vectors encompassing coding regions of the gene. Gene repression comprises introducing a gene ablation sequence or a repressor of the gene to the cell.

Cells which have one or more genes ablated may also be used. For example, a metastatic suppressor gene may be ablated to prevent inhibition to metastases. A useful gene for ablation is a gene capable of affecting the phenotype and behavior of a cell or tumor. For example, with prostate tumors, suitable genes include both well known genes and genes isolated by the methods of the invention such as for example, *TGF-β1, Cyclin D1, p21, p34, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb and α actinin 3.* Genetic ablation (gene knockout) refers to a process of silencing the expression of a particular gene in a cell. The silencing process may include, for example, gene targeting or antisense blocking. Gene targeting refers to a process of introducing a nucleic acid construct into a cell to specifically recombine with a target gene. The nucleic acid construct inactivates the targeted gene. Inactivation may be by introduction of termination codons into a coding region or introduction of a repression site into a regulatory sequence. Antisense blocking refers to the incorporation into a cell of expression sequences which directs the synthesis of antisense RNA to block expression of a target gene. Antisense RNA hybridizes to the mRNA of the target gene to inhibit expression.

The host animal is preferably the same species as the implanted cell. In cases of xenogeneic transplants, the host may be immunocompromised genetically or by treatment with drugs such as immunosuppressants. A host may be immunocompromised genetically by breeding such as with nude mice or severe combined immunodeficient (SCID) mice. A host may also be immunocompromised by chemical or irradiation methods. An additional route to immunocompromise a host is to use transgenic technology to introduce an immunosuppressing gene or to introduce a foreign antigen gene. An immunosuppressing gene is a gene that affects the efficiency of the immune system such as a gene which inhibits the formation of cells of the B cell or T cell lineage. A foreign antigen gene,

when expressed, may cause the host to tolerate the antigens in a xenogeneic transplant and not mount an immune response.

Cells may be implanted into any primary site in a host animal, such as, for example, subcutaneous implantation, intravenous injection, or implantation into the abdominal cardiac, chest, pulmonary, thoracic or peritoneal cavity. Using techniques known to those of ordinary skill in the art, cells can be placed on or in nearly any organ or tissue. Reasons for choosing a site include ease of implant, proximity of similar tissue type, immunoprivileged position and ease of inspection. Metastasises migrate from the primary site to one or more secondary sites such as, for example, the lung, kidney, liver, lymph nodes, brain, testis, bone, spleen, ovaries or mammary. Preferred sites include the renal capsule, the testes, the prostate and the ovaries.

To avoid histocompatibility problems, the implant may be placed into a histocompatible host animal. Such problems are generally avoided if the host animal are syngeneic. Alternatively, a non-histocompatible host may be used if the host can be made immunotolerant. Hosts may also be transgenic or immunocompromised animals or genetically matched to the mammalian cells to be introduced. Immunocompromised animals may be derived from established mouse lines such as nude mice or severe combined immune deficiency (SCID) mice, or by treatments such as radiation, chemical, pharmaceutical or genetic targeting. Sufficiently immunosuppressed animals can be made tolerant to xenogeneic transplants.

After implantation the host animal is maintained under normal conditions to develop metastases. Alternatively, the host animal may be subjected to an altered treatment or environmental condition to stimulate or repress metastasis or induce other cellular functions. In metastasis, a sub-population of cells of the implantation site invade and establish one or more secondary colonies in the host animal. The behavior of the implanted cell will depend on the cell type, the transfected sequence and the implantation

location. Typical secondary sites for metastatic colonies include lung, kidney, liver, lymph nodes, brain, testis, spleen, bone, ovary, skin and mammary tissue. Metastatic development times vary from days to weeks even months. Cells with a high metastatic potential tend to progress to metastasis quickly while cells with a low metastatic potential may require very long periods of time that span significant portions of the lifespan of the animal.

The host animal may be analyzed for metastatic development weekly, from one week to 20 weeks to six months, nine months or one year after implantation. For animals with longer lifespans such as sheep, the animal may be inspected yearly from one year on up to ten years for metastatic tumors. Metastases can be detected by examinations such as palpitation, biopsy, imaging, exploratory surgery, CAT scans, autopsy, X-ray and direct observation. In addition, tissue samples may be taken surgically from the host mammal and subjected to histological or other examination for the detection of metastases.

Expressed sequences include mRNA, rRNA, hnRNA, DNA, cDNA and any nucleic acid sequence that is expressed in the cell. These sequences may be amplified by *in situ* techniques or by purification of nucleic acid from collected cells. Expressed sequences may be obtained by extracting nucleic acids from cells before implantation, at the primary site or at the secondary site. Cells collected at these sites may optionally be cultured for a time before nucleic acid extraction. The effects of treatment with gene expression modifying agents or environmental conditions can be ascertained by collecting cells before and after treatment. Treatment may be applied to the cells while the cells are in the host mammal or after the cells are excised and in culture. Nucleic acid are collected from cells using techniques that are well known to those of ordinary skill in the art.

Expressed sequences may be used directly for polymerase chain reaction (PCR) analysis using, for example, the technique of reverse

transcriptase polymerase chain reaction (RT-PCR). Alternatively, RNA may be enriched for mRNA using a poly-A RNA enrichment method. Numerous poly-A RNA enrichment methods exist and are commercially available. Techniques used for poly-A RNA enrichment include oligo-dT columns, oligo-dT magnetic beads, and oligo-dT cellulose. RNA may be further processed into cDNA before analysis by reverse transcription using reverse transcriptase. The cells or the extracted nucleic acid may be preserved, such as by freezing, and analyzed at a later time.

Differential display polymerase chain reactions (DD-PCR) are performed on the expressed sequences using two variable primers which may contain the same or entirely different sequences or an anchor primer and a variable primer. If an anchor primer is used, one anchor primer and one variable primer create a single or a single set of reaction products for each reaction. A complete profile may include 25 or more different PCR reactions per sample wherein each PCR reaction is performed with the same anchor primer and a different variable primer. DD-PCR may also be performed using anchor and variable primers which contain the same sequence. Whether a particular reaction is used depends on whether a difference exists between the products of two PCR reactions using the same primers. When a significant difference exists between the expression sequences amplified, one pair of PCR reactions may be sufficient and informative.

Anchor primers are preferably oligonucleotides with a poly-T sequence at the 5' -terminas and a dinucleotide selected from the group consisting of AA, AG, AC, AT, GA, GG, GC, GT, CA, CG, CC and CT at the 3'-terminas. For example, the sequence may be 5'-TTTTTTAA-3' or 5'-TTTTTTAG-3'. The length of the poly-T sequence is typically between about 5 to about 30 bases in length and preferably between about 10 to about 20 nucleotides long. The total length of the anchor primer can vary greatly for each experiment but is preferably between about 7 to about 32 and more

preferably between about 12 and about 22. Differential diagnostic polymerase chain reaction may also be performed using an anchor primer of any sequence and a length between about 5 to about 30, preferably between about 5 to about 20 and more preferably between about 7 to about 12 bases.

The variable primer may comprise a random sequence, or a specific sequence such as, for example, a sequence of SEQ ID NO. 1 to SEQ ID NO. 24. Variable primers preferably are oligonucleotides with a length between about 5 to about 30, preferably between about 5 to about 20, and more preferably between about 7 to about 12 bases in length.

To enhance detection of the PCR product, the anchor primer or the variable primer, or both, may comprise a detectable moiety. Examples of detectable moieties include radioactive moieties, phosphorescent moieties, magnetic moieties, luminescent moieties, conjugatable moieties or other detectable moiety. A plurality of detectable moieties may be used to enhance detection or to simplify data analysis. Other detectable moieties include conjugatable moieties and molecules which can bind specifically to other molecules which are themselves detectable. Examples of conjugatable moieties include avidin, streptavidin, biotin, antibody, antigen, cell adhesion molecules and other molecules with similar activities. Detectable moieties are preferably labeled nucleotides. A nucleotide may be any natural or synthetic nucleotide or nucleotide analog capable of incorporation into an elongation reaction in a polymerase chain reaction. Labeled nucleotides include nucleotide triphosphates labeled with one or more radioactive atoms such as ^{32}P, ^{33}P, ^{3}H, ^{14}C and ^{35}S.

Products of DD-PCR reactions are compared to detect the metastatic sequence. Comparisons can be performed between expressed sequences from cells at secondary sites with cells at any stage in the method including untreated mammalian cells, transfected or treated mammalian cells, implanted cells or cells obtained from the primary site in the host

animal. DD-PCR products may be analyzed by any method which reliably compares the products of two polymerase chain reactions. Typical analytical methods used for this purpose include polyacrylamide gel electrophoresis, capillary electrophoresis and high pressure liquid chromatography (HPLC). Product produced from DD-PCR may be analyzed in double-stranded or single-stranded forms. When the products of the DD-PCR reaction are labeled the sizes and distribution of the products may be monitored and analyzed by following the labels using a radiation monitor or by autoradiography. For example, DD-PCR performed in the presence of radioactive primers or nucleotide triphosphates, can be analyzed by gel electrophoresis, by capillary electrophoresis, or by HPLC. Products are easily monitored by the presence of radioactivity.

Another method for analyzing and isolating metastatic sequences is to sequence the amplified nucleic acid sequences. Sequencing may be performed using standard methods well known to those of ordinary skill in the art. The resulting sequence may be compared to a sequence database created or well-known, such as Genbank, for identification or for locating homologs. The sequencing information may be used to calculate the physical characteristics of the nucleic acids such as melting temperature and secondary structure. The primary sequence and the physical characteristic may be used to synthesize optimal nucleic acid probes for the detection or staging of metastasis or conditions that are predictive of the presence or absence of the metastatic condition.

Another embodiment of the invention is directed to a methods for identifying a metastatic sequence. A mammalian cell is pretreated with a metastatic agent to form a population of cells predisposed to metastasize. The treated cells are introduced into a host mammal at a primary site. The host animal is maintained for a period of time sufficient to develop a metastasis at a secondary site. Expressed sequences of cells at the primary site and cells at the secondary site are treated with a genotoxic agent or

subjected to genotoxic conditions. Expressed sequences of the treated cells are amplified by differential display polymerase chain reaction and compared with untreated cells from any previous step to identify the metastasis sequence.

The metastatic agent may be a chemical compound, a nucleic acid or a protein that alters the metastatic potential of a cell or relates to or is associated with the metastatic process. Chemical compounds include retinoids such as 4-hydroxyphenyl (4HP). Other agents include the proteins TGF-β1, Cyclin D1, p21, p34, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb or α-actinin 3, or their respective genes. The metastatic agent may be a metastatic stimulant or a metastatic suppressant. Metastatic stimulants may be used to enhance the sensitivity of the metastasis sequence detection method. Conversely metastatic suppressants may be used to decrease the sensitivity of the method enabling the selective identification of potent metastatis sequences or sequences specific to a particular tissue type or metastatic disorder. Treatment may comprise direct contact with the metastatic agent or incubation for a period of time. Metastatic agents enhance the metastatic potential of the implanted cells and increase the sensitivity and the speed of the overall method.

The cells at the primary site and the metastatic cells at the secondary site may be treated with a genotoxic agent *in vivo* or *in vitro*. *In vivo* treatment may comprise injecting genotoxic agents directly into the host mammal or specifically applying the agent with, for example, topical formulations. The cells at the primary site and the secondary site may also be isolated from the host animal and treated with the genotoxic agent in culture. Genotoxic agents are chemical compounds, nucleic acids or proteings that alter gene expression by effecting the nucleic acid genome directly by, for example, chemical modification, or indirectly by, for example, altering components associated with gene expression. Such agents

include, for example, benzanthracene (BA), dimethyl benzanthracene (DMBA) and 5-azacytidine, and may include metastatic agents as well. In addition to or in place of genetoxic agents, the cells may be treated to hypoxic conditions or radiation to alter gene expression. Metastatic sequences identified in these methods may be specific for particular genotoxic agents or conditions.

Another embodiment of the invention is directed to the use of a host animal with an altered genotypic or phenotypic predisposition for metastases. A host animal may be screened for endogenous expression of metastases gene. Examples of metastatic sequences which may be screened for include sequences isolated by the method of the invention, such as, for example, the sequences listed in Figure 12 and Figure 13. Particularly useful metastatic sequences include *TGF-β*. A host animal with reduced levels of a metastatic gene product may be used to isolate novel metastatic genes. Host animals may be screened for reduced levels of metastatic gene expression. In addition, transgenic technology may be use to ablate a metastatic gene in the germline of a host animal.

Another embodiment of the invention is directed to analysis of a cell line before their use as a starting material to isolate metastatic genes in a particular pathway. Analysis is useful in identifying cells, and consequently sequences specific to these cells, which are particularly susceptible or resistant to metastatic transformation. For example, a cell highly predisposed to metastasis may be especially sensitive for detecting metastatic genes. Conversely, a cell showing high resistance to metastasis can be used to isolate especially potent metastatic sequences. One method to analyze susceptibility to metastasis is to determine the cellular response to growth factors or growth inhibitors. Briefly, a control population and a test population of cells are exposed to a growth factor or a growth inhibitor and the cellular response (e.g. proliferation, metabolism) recorded. Cells showing abnormal responses to the growth factor or growth inhibitor may be

used as the starting material for metastatic gene isolation. Cellular response include changes in the rate of cellular division (e.g. thymidine uptake), changes in the expression of RNA or proteins, changes in cellular localization or modification patterns of RNA or proteins, and changes in the rate of uptake, release or metabolism of nutrients.

Especially potent or weak metastatic genes may be detected by treating and analyzing the metastatic potential of different cells and selecting a suitable cell type as the starting material. For example, cells may be treated with *myc, ras, p53* or combinations thereof and analyzed for *cyclin D1* expression which is shown to correlates with metastasis. Figure 2 shows the *in situ* analysis of *cyclin D1* in primary MPR tumors (Figure 2A) and in metastatic deposits from the lung of the same animal (Figure 2B). The gene expression pattern of *cyclin D1* in MPR correlates with that of human prostate tumors (Figure 3) analyzed with stains specific for *cyclin D1* expression. Normal human tissue shows no *cyclin D1* expression or staining (Figure 3A). Moderately differentiated prostate cancers with dispersed (Figure 3B) or focal positively staining (Figure 3C) show moderate staining. Advanced poorly differentiated prostate cancer show strong nuclear as well as cytoplasmic staining (Figure 3D) implying strong expression of *cyclin D1*. After treatment with *myc*, *ras* or *p53*, *cyclin D1* expression shows correlation with the metastatic potential of the cell. Thus, *cyclin D1* expressing cells are a source of cells with high metastatic potential. Conversely, cells with low *cyclin D1* expression are a source of potentially metastatically resistant cells.

This method may be adjusted for the isolation of metastatic sequences expressed along a particular developmental or differentiation pathway by combining the various treatment and analytical techniques. This approach is schematically represented in Figure 4. For example, a mammalian cell may be genetically ablated for *TGF-β1, Cyclin D1, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme*

E1, nmb, α actinin 3, or *p34.* The genetically altered cell is used in a *in vivo* mouse prostate reconstitution (MPR) model. Metastatic and nonmetastatic cells isolated from the MPR may be analyzed directly or after induction with an agent such as the *TGF-β* gene or its product. Analysis involves the use of differential display polymerase chain reaction to identify differentially expressed bands. Sequences identified may be used for subsequent ablation, transformation or differential analysis.

Genetic ablation (gene knockout) may be performed after a cell is selected or by selecting a cell comprising a genotype with the proper genetic ablation. Cells already comprising gene ablation may be acquired from a cell depository, from other laboratories or from a transgenic animal. As transgenic animals comprise genetically ablated genes in every cell, any tissue from a transgenic animal may be used as the starting material.

The effects of oncogenes are at least additive and often synergistic. Thus, dominant oncogenes may be transfected together or multiple recessive oncogenes ablated together for a stronger effect. Furthermore, both methods may be combined and dominant oncogene transfection may be accompanied by recessive oncogene ablation.

The function of the metastatic sequence may be determined by the differential expression pattern. For example, a dominate metastatic gene will be present in a metastatic cell while a recessive metastatic gene is present in a non-metastatic cell. Metastatic sequences may be detected as bands which are present in the DD-PCR of metastases isolated in secondary sites and absent from DD-PCR products of primary cells. These sequences may be dominant metastatic genes whose expression is directly responsible for metastases, or they may be metastasis associated genes whose expression correlates with metastasis. Either are useful for therapy and diagnosis. Conversely, DD-PCR bands which are present in primary site tumors, but absent in secondary metastatic sites, may be dominant metastasis suppression genes. Dominant metastasis suppression genes comprise genes

whose expression suppresses metastasis while nonmetastatic genes comprise genes whose expression correlates with non-metastatic tissue. Genes which are highly correlative with either the metastatic phenotype or the non-metastatic phenotype may be isolated. Isolation can be performed by cutting the appropriate nucleic acid in the band of a polyacrylamide gel or by collecting the appropriate fraction in an HPLC or capillary electrophoresis. The nucleic acid may be cloned into a plasmid vector, and sequenced, or synthetically prepared.

Another embodiment of the invention is directed to a method for identifying sequences in a metastatic pathway which are responsive or unresponsive to extracellular signals. Such sequences may be used in therapy and diagnosis of metastatic disorders. Implanted cells or cells from a primary site and cells from a secondary site are treated with extracellular signals. RNA sequences from the treated cells are compared with RNA sequences of the untreated cells (Figure 5B). Treated cells and untreated cells may be derived from a short term or long term *in vitro* culture of primary tumor and malignant tumors. Alternatively, a part of a primary tumor and a part of a malignant tumor may be collected before the animal is treated with an extracellular cytokine or other factor. Long term cultures, or cell lines of primary and malignant cells may also be used as recipients of extracellular growth signal treatment. Suitable signals for each experiment will depend on the cell type. Generally, growth factors, lymphokines, inhibitory factors, migratory factors or hormones may be used. Factors previously isolated by commercial or methods of the invention and factors associated with or causative or suppressive of metastasis are preferred. Thus, transforming growth factor β1 (TFG-β1) may be used to treat cells before DD-PCR analysis. Proteins encoded by the genes isolated by this method are especially useful for the treatment of cells for the isolation of additional sequences. The identification of one sequence responsive to the

extracellular signal pathway allows for identification of additional genes upstream and downstream from that sequence.

Another embodiment of the invention is directed to metastatic sequences identified by the methods of the invention. Metastatic sequences are sequences associated with the presence or absence of a metastasis or related to the metastatic processor can be used in the therapeutic treatment of metastasis. Metastatic-related sequences include dominant metastatic sequences, recessive metastatic sequences, metastasis associated sequences, dominant oncogenes, recessive oncogenes and cell cycle genes. These genes encode for example, proteins involved in cell cycle, signal processing, DNA replication, growth regulation, inter and intra cellular signaling transcription control and translation control. Isolated sequences are useful in the treatment and for the detection of metastatic and other disorders. Disorders which may be treated comprise diseases involving proteins and sequences which are isolated by interaction with the sequences and proteins isolated by the method of the invention. Both malignant or nonmalignant disorders may be treated. Non malignant disorders include hyperplasia, dysplasia and hypertrophy. Examples of nonmalignant disorders include benign enlargement of the prostate, nodular hyperplasia, and benign prostatic hypertrophy.

Treatment may involve gene replacement, gene targeting, antisense inhibition, gene expression or gene suppression. Gene replacement involves replacing a copy of a defective gene with another copy by homologous recombination. Gene targeting involves the disruption of a cellular copy of a gene by homologous recombination. Antisense inhibition exploits the specificity of hybridization reactions between two complementary nucleic acid chains to suppress gene expression. Cloned genes can be engineered to express RNA from only one or the other DNA strands. The resultant RNA hybridizes to the sense RNA and inhibit gene expression. Gene expression and gene suppression involve the introduction

of genes whose expression actively inhibits neoplastic transformation and metastasis.

Another embodiment of the invention is directed to nucleic acids which comprise a sequence identified by the methods of the invention. The nucleic acid may be DNA, RNA or PNA and may be used as a diagnostic tool in the treatment of neoplastic disorders and malignant tumors. The nucleic acids may comprise additional sequences such as promoters, for expression of a sense or antisense message, recombination sequences for gene targeting, selectable markers for transfections, or replication origins for passage in a prokaryotic or eukaryotic host such as animal cells, bacteria or yeast.

Another embodiment of the invention is directed to nucleic acids which comprise sequences identified by the method of the invention such as, for example, the caveolin, ABP280 (actin binding protein 280), the lysyl oxidase gene, and the nmb gene (clone 29), and other sequences listed in Figure 12 and Figure 13. Nucleic acids comprising a sequence corresponding to these genes may be used in treatment or diagnosis and in diagnostic kits for screening biological samples for the presence or absence of metastasis or metastatic potential. Treatment may involve using the sequences in gene therapy, including gene ablation, gene expression and antisense suppression. Diagnosis may involve genotypic analysis of samples to determine the existence and expression levels of the expressed sequences.

Another embodiment of the invention is directed to the use of caveolin gene and protein in the isolation of oncogenes and in the treatment of neoplastic disorders such as, for example, prostate cancer. Caveolin is an integral membrane protein and a principal component of caveolae. Caveolae are small invaginations at or near the plasma membrane of most smooth muscle cells and may function as a component of specific signal transduction pathways. Surprisingly, caveolin expression increases in metastatic human prostate cells as compared to human primary prostate tumors.

As caveolin expression correlates with metastasis, application of biological technologies designed to block the activity of caveolin or the function of caveolae may have therapeutic benefits for the treatment of neoplastic disorders such as human prostate tumors. Specific treatment approaches using caveolin may include the delivery of antisense or dominant negative caveolin sequences using expression or viral vectors; as well as the use of specific anti-caveolin antibodies. Additional approaches could also target the cavoeolae, but are not specifically based on caveolin function. Additional protein and non-protein components of caveolae could also be targeted for abrogation or the local or systemic administration of nutritional or biological agent may also be used. For example, caveolae are extremely rich in cholesterol and disruption or depletion of this molecule may alter the function of caveolae.

Another embodiment of the invention is directed to methods for treating a neoplastic disorder comprising administering a pharmaceutically effective amount of composition containing a nucleic acid having a sequence identified according to the methods of this invention, its expression product or fragments of either. The nucleic acid may be in the form of a sense or antisense single-stranded or double-stranded nucleic acid. The composition may be combined with a pharmaceutically acceptable carrier such as water, alcohols, salts, oils, fatty acids, saccharides, polysaccharides administered by injection, pulmonary absorption, topical application or delayed release. More than one carrier may be used together to create a pharmaceutical with desirable properties.

Another embodiment of the invention is directed to a kit or diagnostic acid for screening biological samples for detection of metastasis, neoplasia or kits comprise sequences isolated according to the methods of the invention and reagents and materials useful in such kits, such as, for example, buffers, salts, preservatives, and carriers, all of which are well known to those of ordinary skill in the art. Kits are useful for the analysis

of tissues to screen those for the determination of normal, nonmalignant neoplastic or malignant cells. Kits may comprise additional reagents useful for the extraction of nucleic acids from a tissue sample. Reagents for analyzing the nucleic acid extracted from a tissue sample such as polymerase chain reaction reagents and Southern blots reagents may also be included.

The following experiments are offered to illustrate embodiments of the invention and should not be viewed as limiting the scope of the invention.

Examples

Example 1 Production of Mouse Prostate Reconstitution Tumors and Metastasis.

Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mice embryos. Each isolated UGS was digested with 1% trypsin for three hours at 4°C. The trypsin was inactivated by the addition of fetal calf serum. UGS cells were digested with 0.125% collagenase for 1.5 hours, counted and mixed at the appropriate cell rations prior to infection with retrovirus in the presence of polybrene. Retroviruses used include Zipras/myc-9. Control experiments were performed using BAGα virus. After a two-hour infection, the infected cells were centrifuged and individual reconstitutions containing 1.5×10^6 cells produced by resuspending the cells in rat tail collagen at a density of 6.0×10^7 cells per ml. Aliquots of the infected UGS cells were placed in (DME) with 10% fetal calf serum overnight at 37°C, 5% CO_2.

The next morning each cell/collagen reconstitution was implanted under the renal capsule of an adult male +/+ animal. Reconstitutions were harvested from the mice five weeks later when they showed signs of obvious distress from the tumor burden. Metastasized tumors were isolated from the same mice at sites outside the renal capsule.

Isolated tumors and metastasises were either stored in liquid nitrogen or in preservatives such as 10% buffered formalin.

Cell lines were derived from fresh tumors by mincing a small portion of the primary metastatic or nonmetastatic tumor and placing each in explant culture in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal calf serum. Cells which grow from each explant were propagated in DMEM and 10% fetal calf serum.

For histological analysis, a portion of a fresh tumor was fixed in 10% buffered formalin and embedded in paraffin for sectioning and staining with hematoxylin and eosin (H&E) or immunohistochemical staining. Immunohistochemical localization of cytokeratins was detected using polyclonal cytokeratin antiserum A575 (Dake Co.; Carpinteria, CA) and Vectastain ABC kit (Vector Laboratories; Burlingame, CA).

Example 2 Isolation of C-DNA for DD-PCR.

Total cellular RNA was isolated by ultracentrifugation through cesium chloride. Briefly, up to one gram of cells from culture, tumors or organs was placed into 4 ml of ice-cold GIT buffer (4M guanidine isothiocyanate, 0.025 M sodium acetate, 0.1 M β-mercaptoethanol) and homogenized in a tissue homogenizer (Polytron or equivalent). The homogenate was carefully layered over 4 ml of 5.7 M CsCl, 0.024 M sodium acetate (1.8 g CsCl per ml) in a centrifuge tube. The layers were centrifuged at 35,000 RPM for 18 hours in a SW50.1 rotor. DNA was collected from the interface between the cushion and the supernatant, diluted two folds with water, added to 2.5 volumes of ethanol and spooled out on a glass rod. RNA that formed a pellet on the bottom of the CsCl layer was resuspended, and once extracted with an equal volume of phenol:chloroform (1:1), twice with chloroform and precipitated with ethanol and resuspended in diethylpyrocarbonate treated water. The concentration of DNA and RNA were be determined by absorption at 260 nanometers.

Example 3 Differential Display Polymerase Chain Reaction.

mRNA isolated from primary tumors or metastasis was reverse transcribed with one of the primers and subjected to DD-PCR using the same primer as both the forward and reverse primer. A set of 24 primers comprising short oligonucleotides were used for both the reverse transcription of mRNA into c-DNA and for differential display polymerase chain reaction. The sequence of the primers used are shown in Table 1.

Table 1

Primer No.	Sequence	Sequence number
1	5'-TGACAATCG-3'	(SEQ. ID. NO. 1)
2	5'-AGCTAAGGTC-3'	(SEQ. ID. NO. 2)
3	5'-TCTGCGATCC-3"	(SEQ. ID. NO. 3)
4	5'-ATACCGTTGC-3'	(SEQ. ID. NO. 4)
5	5'-TACGAAGGTG-3'	(SEQ. ID. NO. 5)
6	5'-TGGATTGGTC-3'	(SEQ. ID. NO. 6)
7	5'-CTTTCTACCC-3'	(SEQ. ID. NO. 7)
8	5'-GGAACCAATC-3'	(SEQ. ID. NO. 8)
9	5'-TGGTAAAGGG-3'	(SEQ. ID. NO. 9)
10	5'-TCGGTCATAG-3'	(SEQ. ID. NO. 10)
11	5'-CTGCTTGATG-3'	(SEQ. ID. NO. 11)
12	5'-GATCAAGTCC-3'	(SEQ. ID. NO. 12)
13	5'-GATCCAGTAC-3'	(SEQ. ID. NO. 13)
14	5'-GATCACGTAC-3'	(SEQ. ID. NO. 14)
15	5'-GATCTGACAC-3'	(SEQ. ID. NO. 15)
16	5'-TTAGCACCTC-3'	(SEQ. ID. NO. 16)
17	5'-ACCTGCATGC-3'	(SEQ. ID. NO. 17)
18	5'-GCTATACTGC-3'	(SEQ. ID. NO. 18)
19	5'-AGTTGCCAGG-3'	(SEQ. ID. NO. 19)

20	5'-AAGCCGTGTC-3'	(SEQ. ID. NO. 20)
21	5'-TCAACGCTCA-3'	(SEQ. ID. NO. 21)
22	5'-TGTTCGAATC-3'	(SEQ. ID. NO. 22)
23	5'-CGAGTCAGAC-3'	(SEQ. ID. NO. 23)
24	5'-TATGAGTCCG-3'	(SEQ. ID. NO. 24)

PCR was performed using standard conditions with 40 cycles of denaturation at 94°C for 40 seconds, annealing at 40°C for 2 minutes, and elongation at 72°C for 35 seconds. After PCR, the products were analyzed with non-denaturing polyacrylamide gel electrophoresis (PAGE) at 12 watts for 15 hours. Bands which differed between test and control samples were eluted from the gel, subjected to reamplification by PCR and cloned. Polyacrylamide gel electrophoresis of DD-PCRs, and the accompanying RNA blot analysis showing the isolation of sequences with substantial similarity to nmb and TGF-β is shown in Figure 6 and Figure 7 respectively. Additional sequences isolated by this method show substantial similarity to lysyl oxidase, actin binding protein, ubiquitin activating enzyme E1, α-actinin, and P34 ribosomal binding protein sequence (Figure 8). Differential expression of caveolin was demonstrated by DD-PCR followed by PAGE (Figure 9).

Example 4 *p53* Allelotype Determination.

The *p53* allelotype of a cell sample was determined by PCR. Briefly, nucleic acid is extracted from a tissue sample or a cell culture sample. An aliquot of nucleic acids in placed in 45 µl aliquot of a master mix which contained a final concentration of 0.2 mM of each dATP, dTTP, dGTP, dCTP, 1.5 mM $MgCl_2$, 0.5 unit Taq polymerase, 0.05 µM of each of two primers set specific for the normal wildtype allele of *p53* (5'-GTGTTTCATTAGTTCCCCACCTTGAC-3', SEQ. ID NO. 25; 5'-

AGAGCAAGAATAAGTCAGAAGCCG-3', SEQ. ID NO. 26). A control set of primers specific for the fibroblast growth factor-7 gene was used to monitor the polymerase chain reaction experiment (5'-ACAGACCGTGCTTCCACCTCGTC-3', SEQ. ID NO. 27; 5'-CCTCATCTCCTGGGTCCCTTTCA-3', SEQ. ID NO.28). One µl of the reaction from the first round of PCR was used as the starting material for a second round of PCR using a second set of wildtype *p53* specific primer (5'-GTCCGCGCCATGGCCATATA-3', SEQ. ID NO. 29; 5'-ATGGGAGGCTGCCAGTCCTAACCC-3', SEQ. ID NO. 30). This second round of PCR was also monitored using a control set of primers specific for the fibroblast growth factor-7 (5'-ACAGACCGTGCTTCCACCTCGTC-3', SEQ. ID NO 27; 5'-CCTCATCTCCTGGGTCCCTTTCA-3', SEQ. ID NO 28).

After PCR the products were analyzed with non-denaturing polyacrylamide gel electrophoresis (PAGE) at 12 watts for 15 hours. Bands which differed between test and control were eluted from the gel, subjected to reamplification by PCR and cloned.

Example 5 Induction of cell lines with *TGF-β1* Influence Cellular Gene Expression.

1481-PA cells were grown overnight in DME supplemented with 10% fetal calf serum overnight at 37°C, and 5% CO_2. Induction was performed by treatment with *TGF-β1* at a concentration of 2 nanograms per ml. The treated cells were returned to the incubator and cultured for 12 hours. After induction, cells were washed in phosphate buffered saline and harvested and concentrated by centrifugation.

RNA was extracted from treated and untreated cells and subjected to DD-PCR. Differentially expressed bands detected by DD-PCR were cloned and differential expressions were confirmed using RNA blots

(Figure 10). Subsequent cloning and sequencing identified the bands as ABP280 or filamin.

One gene isolated showed differential expression in cells induced by *TGF-β* (Figure 11, clone 29), while a control probe on the same cell line showed no difference in expression levels (Figure 11, GAPDH).

Example 6 Metastatic Sequences Isolated.

Using the methods of Examples 1, 2, 3, 4, and 5, a plurality of metastatic sequences were isolated and sequenced. The expression of the metastatic sequences in primary cells and in metastatic cells were determined using RNA blots. The nucleic acid sequences of other isolated sequences are listed in Figure 12. Sequence analysis and expression analysis was performed on the isolated cloned and the results of these studies are summarized in Figure 13.

Example 7 Caveolin Immunoassay in Human Prostate Cancers.

Primary site human prostate tumors and metastases were isolated and analyzed for caveolin expression by immunoassay. The results of the assay is shown in Table 3. Metastases shows higher levels of caveolin proteins in metastases than in primary tumors. Immunohistology of tissue sections reveals both elevated levels and distinct distribution of caveolin protein in metastatic human prostate when compared to a primary human prostate tumor (Figure 14).

Table 3

Patients	Primary-site	Metastases in lymph node
1	+	++
2	++	+++

3	++	+++
4	++	++
5	+	+
6	++	++
7	++	+++
8	+	+
9	-	-
10	+	+
11	+	+
12	++	++
13	+	+
14	++	+++

Other embodiments and uses of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. The specification and examples should be considered exemplary only with the true scope and spirit of the invention indicated by the following claims.

I Claim:

1. A method for identifying a metastatic sequence comprising the steps of:

a) transfecting an oncogenic sequence into a mammalian cell to form a population of transfected cells;

b) introducing transfected cells to a primary site of a host mammal;

c) maintaining said mammal for a period of time sufficient to develop a metastasis at a secondary site;

d) amplifying expressed sequences of the transfected cells and expressed sequences of the metastasis by differential-display PCR; and

e) comparing the amplified sequences and identifying the metastatic sequence.

2. The method of claim 1 wherein the mammalian cell is transfected by calcium phosphate transfection, viral transduction, lipofection, dextran sulfate transfection or electroporation.

3. The method of claim 1 wherein the oncogenic sequence is a sequence of the gene that expresses the oncoproteins p21, p34, p53, myc, ras or src.

4. The method of claim 1 wherein the oncogenic sequence is a metastatic sequence.

5. The method of claim 4 wherein the metastatic sequence is a sequence of the gene that expresses cyclin D1, caveolin or TGF-β1.

6. The method of claim 1 wherein the oncogenic sequence is a gene ablation sequence specific for the gene that expresses the protein TGF-β1, cyclin D1, p21, p34, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb or α actinin 3.

7. The method of claim 1 wherein the mammalian cell is treated with a metastatic agent that alters gene expression before or after transfection.

8. The method of claim 8 wherein the metastatic agent is benzanthracene (BA), dimethyl benzanthracene (DMBA) or 5-azacytidine.

9. The method of claim 1 wherein the mammalian cell is a primary or established cell line.

10. The method of claim 1 wherein the mammalian cell is derived from urogenital sinus tissue.

11. The method of claim 1 wherein the mammalian cell is a fetal cell.

12. The method of claim 1 wherein the mammalian cell contains a genetically ablated endogenous gene wherein said gene is *TGF-β1, cyclin D1, p21, p34, p53, ras, myc* and homologs thereof.

13. The method of claim 1 wherein the mammalian cell is derived from the same species as the host mammal.

14. The method of claim 1 wherein the mammalian cell and the host mammal are histocompatible.

15. The method of claim 1 wherein the mammalian cell and the host mammal are genetically matched.

16. The method of claim 1 wherein the transfected cell is maintained *in vivo* or *in vitro*.

17. The method of claim 1 wherein a collection of the expressed sequences is obtained from cells at the primary site of the host mammal.

18. The method of claim 1 wherein a collection of the expressed sequences is obtained from a cell line of immortalized transfected cells.

19. The method of claim 1 wherein the transfected cells are introduced to the primary site by subcutaneous implantation.

20. The method of claim 1 wherein the host mammal is a mouse, a rabbit or a primate.

21. The method of claim 1 wherein the host mammal is an syngeneic, xenogeneic, immunocompromised or transgenic host mammal.

22. The method of claim 1 wherein the host mammal has reduced expression of *TGF-β*.

23. The method of claim 1 wherein the primary site is the renal capsule, the prostate or the testis.

24. The method of claim 1 wherein the secondary site is selected from the group of sites consisting of lung, kidney, liver, lymph nodes, brain, bone, testis, spleen, ovaries and mammary.

25. The method of claim 1 wherein differential display PCR is performed with an anchor primer and a variable primer.

26. The method of claim 25 wherein the anchor primer comprises a polythymidine sequence and a dinucleotide sequence connected to a 3'-terminus.

27. The method of claim 26 wherein the polythymidine sequence comprises between about 5 to about 30 thymidines.

28. The method of claim 26 wherein the dinucleotide sequence is selected from the group of sequences consisting of AA, AG, AC, AT, GA, GG, GC, GT, CA, CG, CC and CT.

29. The method of claim 25 wherein the anchor primer or the variable primer comprise a detectable moiety selected from the group consisting of radioactive moieties, phosphorescent moieties, magnetic moieties, luminescent moieties and conjugatable moieties.

30. The method of claim 25 wherein the anchor primer and the variable primer have a common sequence.

31. The method of claim 1 further comprising the step of treating the host mammal with a metastatic agent.

32. The method of claim 31 wherein the metastatic agent is a retinoid.

33. The method of claim 1 wherein identifying comprises determining the nucleotide sequence or expression product of the metastatic sequence.

34. The method of claim 1 wherein the metastatic sequence identified is specifically expressed in metastatic or non-metastatic cells.

35. A metastatic sequence identified by the method of claim 1.

36. The metastatic sequence of claim 35 which is a sequence which encodes *TGF-β1, Cyclin D1, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb, α-actinin 3* or homologs thereof.

37. A method for identifying a metastatic sequence comprising the steps of:

a) pretreating a mammalian cell with a metastatic agent to form a population of cells predisposed to metastasis;

b) introducing the pretreated cells to a primary site of a host mammal;

c) maintaining said mammal for a period of time sufficient to develop a metastasis at a secondary site;

d) treating cells of the primary or secondary sites with a genotoxic agent;

e) amplifying expressed sequences of treated cells by differential-display PCR; and

f) identifying the metastatic sequence.

38. The method of claim 37 wherein the metastatic agent is an oncogenic sequence and the mammalian cell is treated by transfection with the oncogenic sequence.

39. The method of claim 37 wherein the metastatic agent is TGF-β1, Cyclin D1, p21, p34, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb or α-actinin 3, and the mammalian cell is treated by contact with the metastatic agent.

40. The method of claim 37 wherein the genotoxic agent is benzanthracene (BA), dimethyl benzanthracene (DMBA) or 5-azacytidine.

41. The method of claim 37 wherein the metastatic agent and the genotoxic agent are the same.

42. The method of claim 37 wherein the expressed sequences amplified are compared to expressed sequences amplified from mammalian cells before pretreatment to identify the metastatic sequence.

43. The method of claim 37 wherein the expressed sequences amplified are compared to expressed sequences amplified from pretreated cells to identify the metastatic sequence.

44. The method of claim 37 wherein the expressed sequences amplified are compared to expressed sequences amplified from cells obtained from the primary site or cells obtained from the secondary site.

45. A nucleic acid sequence identified by the method of claim 37.

46. A method for identifying a metastatic sequence comprising the steps of:

a) treating a mammalian cell with a metastasizing agent to form a population of treated cells;

b) introducing treated cells to a primary site of a host mammal;

c) maintaining said mammal for a period of time sufficient to develop a metastasis at a secondary site;

d) amplifying RNA sequences of treated cells and RNA sequences of the metastasis by differential-display PCR;

e) comparing the amplified sequences and identifying the metastatic sequence.

47. The method of claim 46 wherein the metastatic agent is a chemical compound, a nucleic acid, a protein or a combination thereof.

48. The method of claim 47 wherein the chemical compound is a benzanthracene, dimethyl benzanthracene, or 5-azacytidine.

49. The method of claim 47 wherein the nucleic acid contains an oncogenic sequence.

50. The method of claim 47 wherein the protein is p53, myc, ras, caveolin or TGF-β1.

51. The method of claim 46 wherein the mammalian cell is transfected with an oncogenic sequence before or after treatment.

52. The method of claim 46 wherein the mammalian cell is a cell line.

53. The method of claim 46 wherein the mammalian cell is derived from lymphatic tissue, hematopoietic cells, reproductive tissues or urogenital sinus tissue.

54. The method of claim 46 wherein the mammalian cell is a fetal cell.

55. The method of claim 46 wherein the mammalian cell is derived from a transgenic animal.

56. The method of claim 46 wherein the primary site is the renal capsule, the prostate or the testis.

57. The method of claim 46 wherein the secondary site is selected from the group of sites consisting of lung, kidney, liver, lymph nodes, brain, bone, testis, spleen, ovaries and mammary.

58. The method of claim 46 wherein differential display PCR is performed using an anchor primer and a variable primer.

59. A metastatic sequence identified by the method of claim 46.

60. A diagnostic kit for screening a biological sample for the presence or absence of metastasis comprising a metastatic sequence identified according to the method of claim 46.

61. A method for treating a metastatic disorder comprising administering a composition containing a therapeutically effective amount of a metastatic sequence or the expression product of said metastatic sequence to a patient wherein said metastatic sequence was identified according to the method of claim 46.

62. The method of claim 61 wherein said metastatic sequence is selected from the group consisting of *TGF-β1, Cyclin D1, p21, p34, p53, lysyl oxidase, caveolin, actin binding protein, ubiquitin activating enzyme E1, nmb, α actinin 3* and homologs thereof.

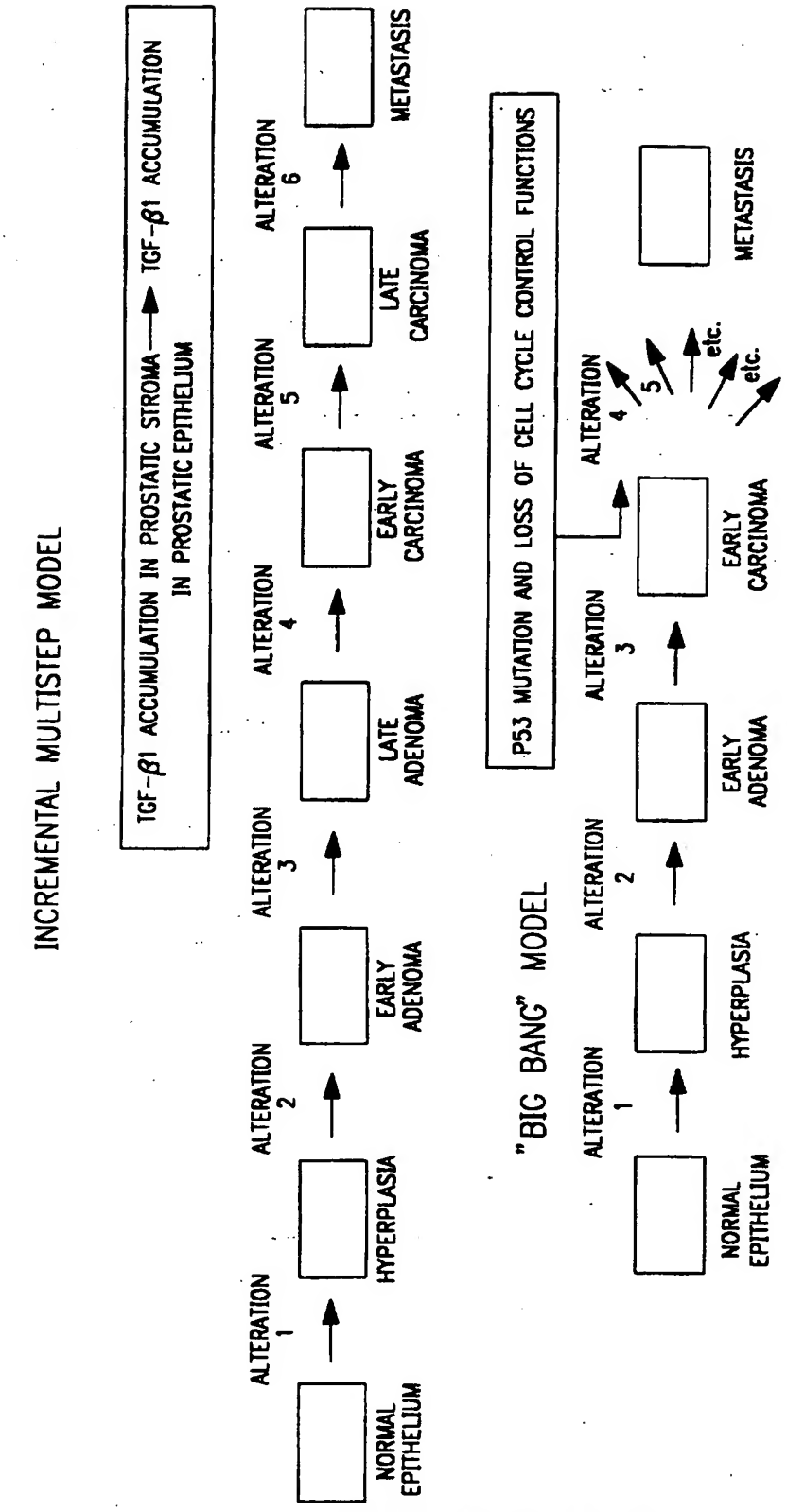
INCREMENTAL MULTISTEP MODEL
TGF-β1 ACCUMULATION IN PROSTATIC STROMA → TGF-β1 ACCUMULATION IN PROSTATIC EPITHELIUM
ALTERATION 1
NORMAL EPITHELIUM
ALTERATION 2
HYPERPLASIA
ALTERATION 3
EARLY ADENOMA
ALTERATION 4
LATE ADENOMA
ALTERATION 5
EARLY CARCINOMA
ALTERATION 6
LATE CARCINOMA
METASTASIS
"BIG BANG" MODEL
P53 MUTATION AND LOSS OF CELL CYCLE CONTROL FUNCTIONS
ALTERATION 1
NORMAL EPITHELIUM
ALTERATION 2
HYPERPLASIA
ALTERATION 3
EARLY ADENOMA
ALTERATION 4
EARLY CARCINOMA
5
etc.
etc.
METASTASIS

FIG. 1

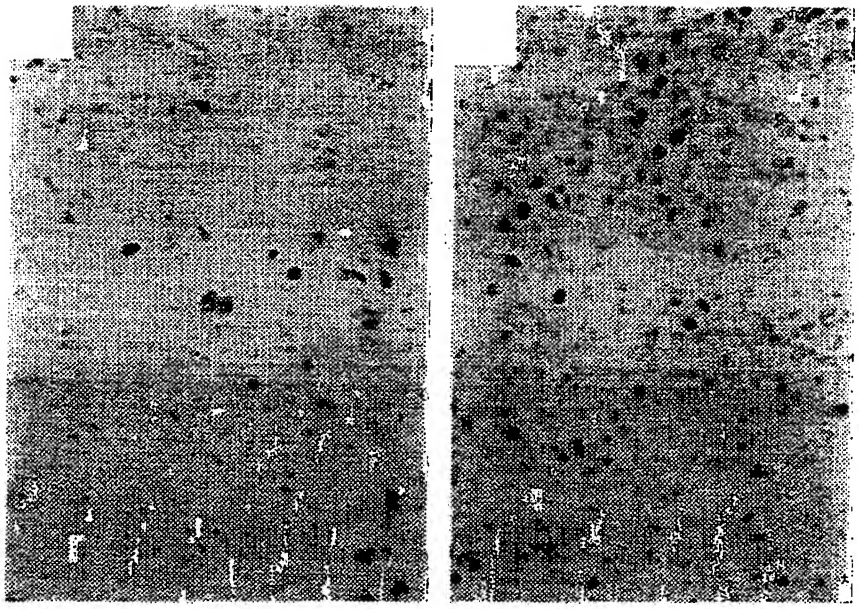

FIG. 2A FIG. 2B

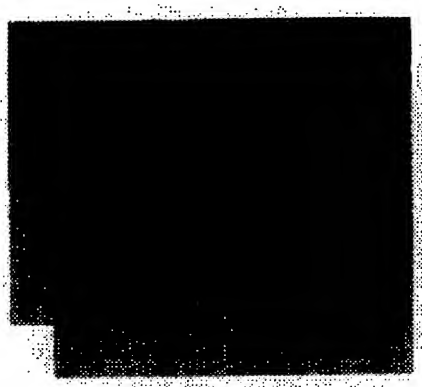

FIG. 3A

FIG. 3B

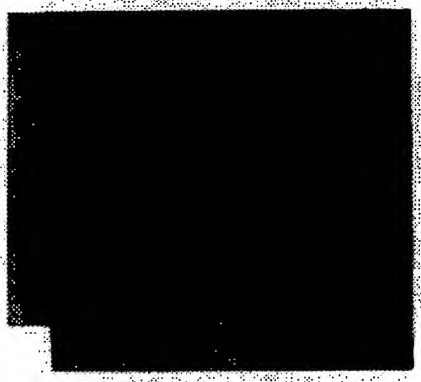

FIG. 3C

FIG. 3D

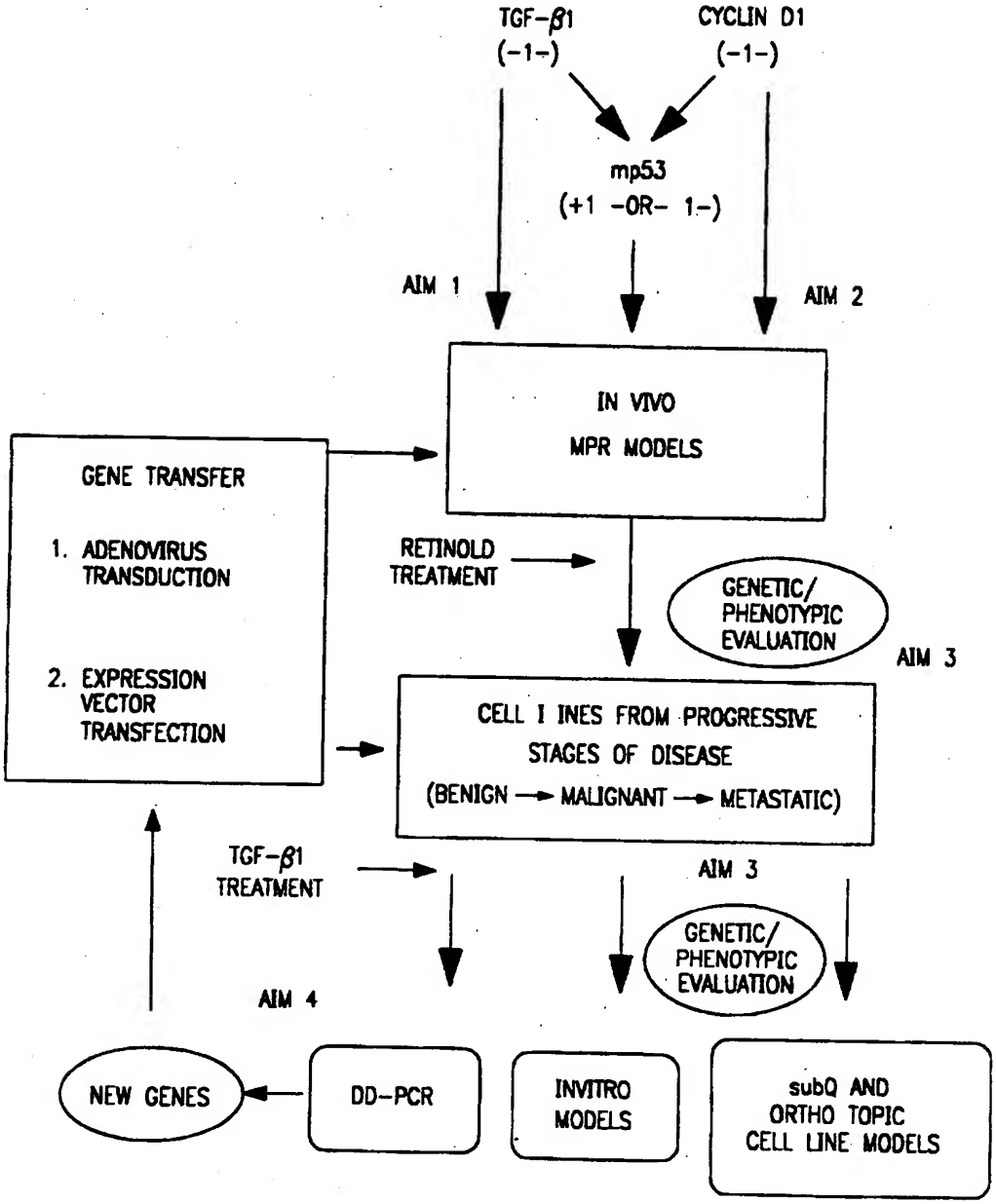
TGF-β1
(-1-)
CYCLIN D1
(-1-)
mp53
(+1 -OR- 1-)
AIM 1
AIM 2
IN VIVO
MPR MODELS
GENE TRANSFER
1. ADENOVIRUS TRANSDUCTION
2. EXPRESSION VECTOR TRANSFECTION
RETINOLD TREATMENT
GENETIC/ PHENOTYPIC EVALUATION
AIM 3
CELL I INES FROM PROGRESSIVE STAGES OF DISEASE
(BENIGN → MALIGNANT → METASTATIC)
TGF-β1 TREATMENT
AIM 3
GENETIC/ PHENOTYPIC EVALUATION
AIM 4
NEW GENES
DD-PCR
INVITRO MODELS
subQ AND ORTHO TOPIC CELL LINE MODELS

FIG. 4

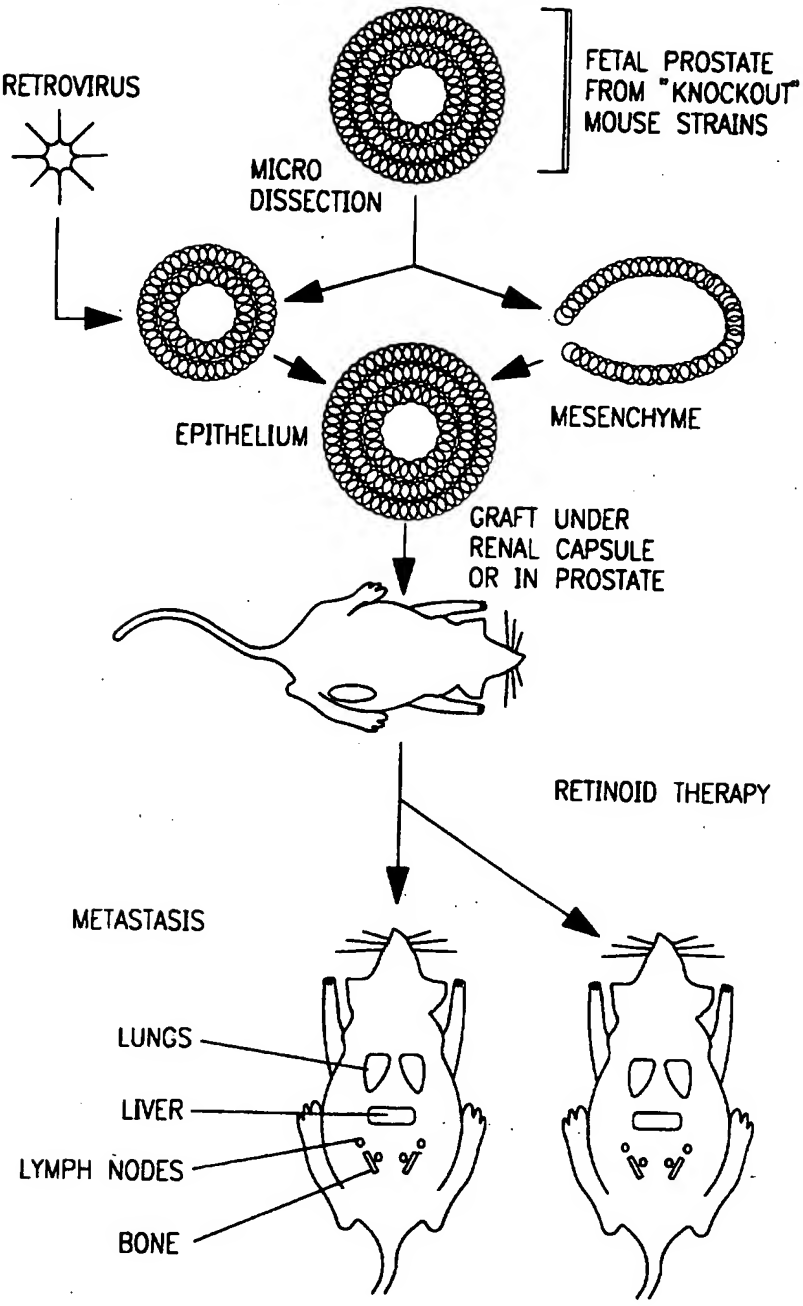
RETROVIRUS
FETAL PROSTATE FROM "KNOCKOUT" MOUSE STRAINS
MICRO DISSECTION
EPITHELIUM
MESENCHYME
GRAFT UNDER RENAL CAPSULE OR IN PROSTATE
RETINOID THERAPY
METASTASIS
LUNGS
LIVER
LYMPH NODES
BONE

FIG. 5A

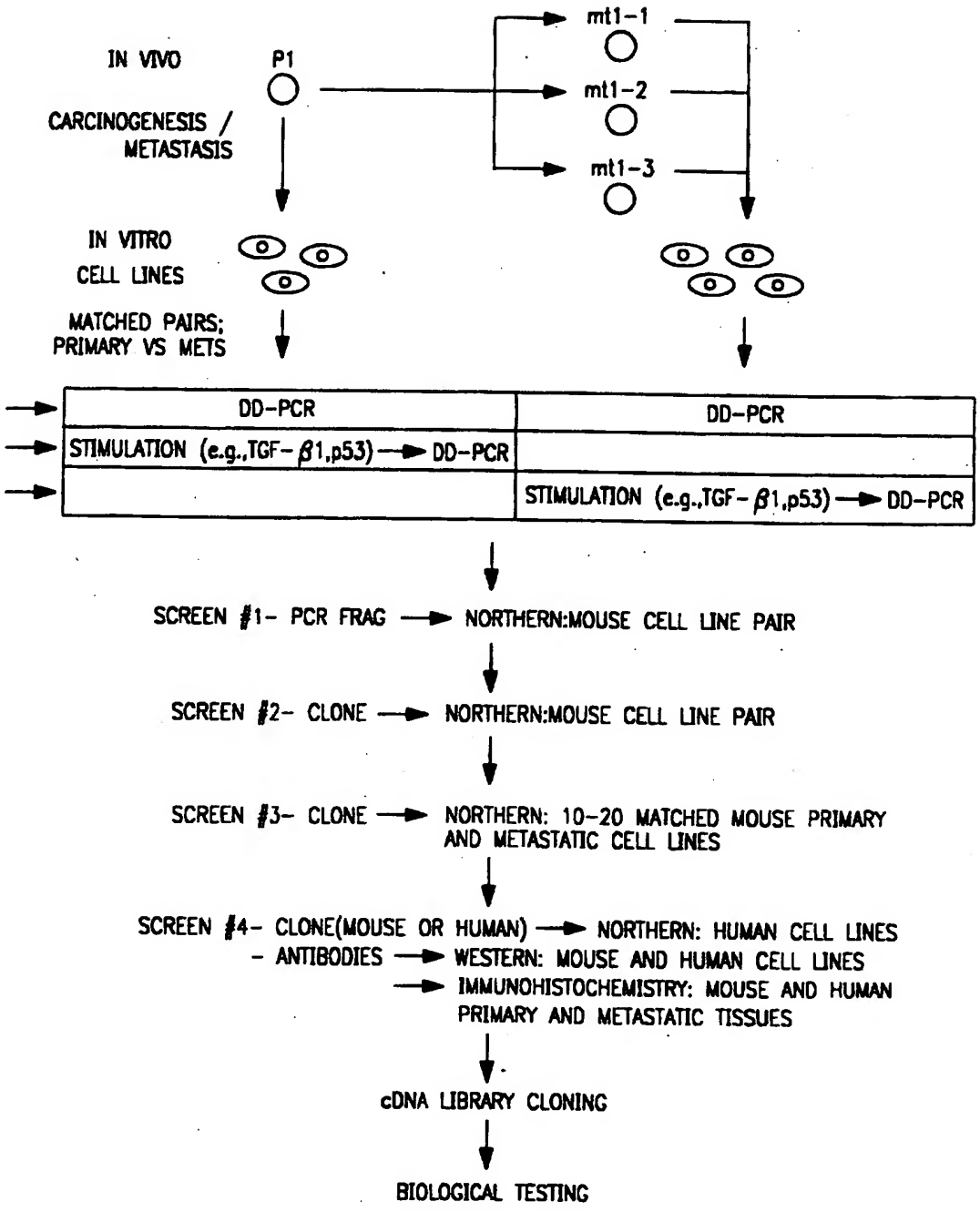
IN VIVO
P1
CARCINOGENESIS / METASTASIS
mt1-1
mt1-2
mt1-3
IN VITRO CELL LINES
MATCHED PAIRS; PRIMARY VS METS
DD-PCR
DD-PCR
STIMULATION (e.g.,TGF-β1,p53) → DD-PCR
STIMULATION (e.g.,TGF-β1,p53) → DD-PCR
SCREEN #1- PCR FRAG → NORTHERN:MOUSE CELL LINE PAIR
SCREEN #2- CLONE → NORTHERN:MOUSE CELL LINE PAIR
SCREEN #3- CLONE → NORTHERN: 10-20 MATCHED MOUSE PRIMARY AND METASTATIC CELL LINES
SCREEN #4- CLONE(MOUSE OR HUMAN) → NORTHERN: HUMAN CELL LINES
- ANTIBODIES → WESTERN: MOUSE AND HUMAN CELL LINES
→ IMMUNOHISTOCHEMISTRY: MOUSE AND HUMAN PRIMARY AND METASTATIC TISSUES
cDNA LIBRARY CLONING
BIOLOGICAL TESTING

FIG. 5B

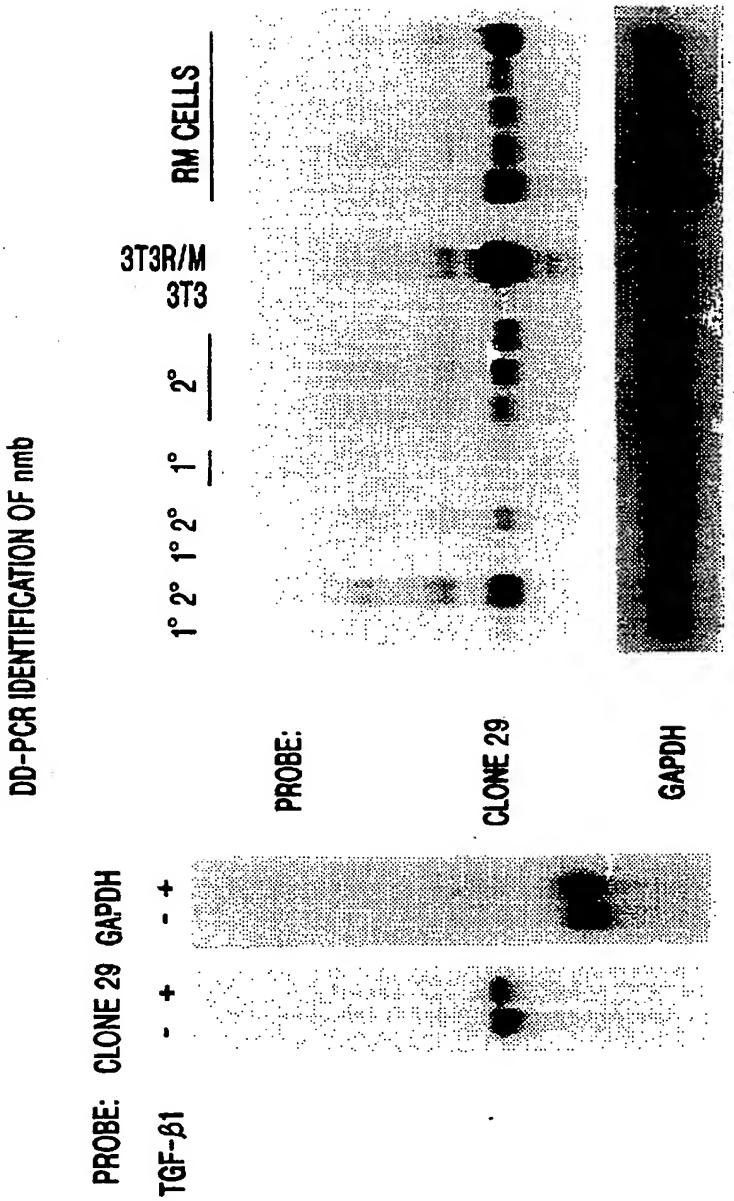
DD-PCR IDENTIFICATION OF nmb
1° 2° 1° 2° 1° 2°
3T3R/M 3T3
RM CELLS
PROBE:
CLONE 29
GAPDH
PROBE: CLONE 29 GAPDH
TGF-β1
- + - +

FIG. 6

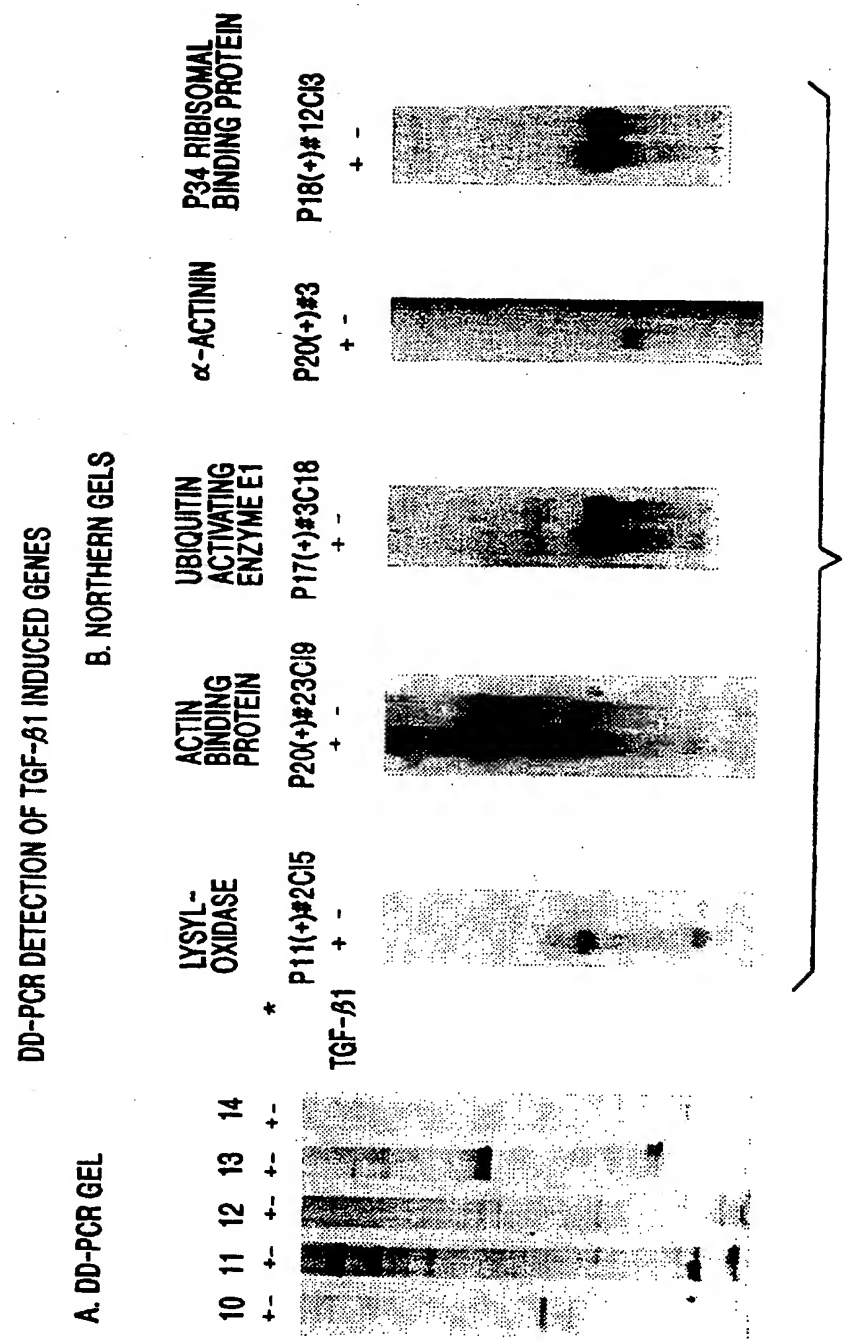
DD-PCR DETECTION OF TGF-β1 INDUCED GENES
A. DD-PCR GEL
B. NORTHERN GELS
PRIMER 10 11 12 13 14
TGF-β1 +- +- +- +- +-
TGF-β1 + -
LYSYL-OXIDASE
P11(+)#2C15
ACTIN BINDING PROTEIN
P20(+)#23C19
UBIQUITIN ACTIVATING ENZYME E1
P17(+)#3C18
α-ACTININ
P20(+)#3
P34 RIBISOMAL BINDING PROTEIN
P18(+)#12C13

FIG. 7A

FIG. 7B

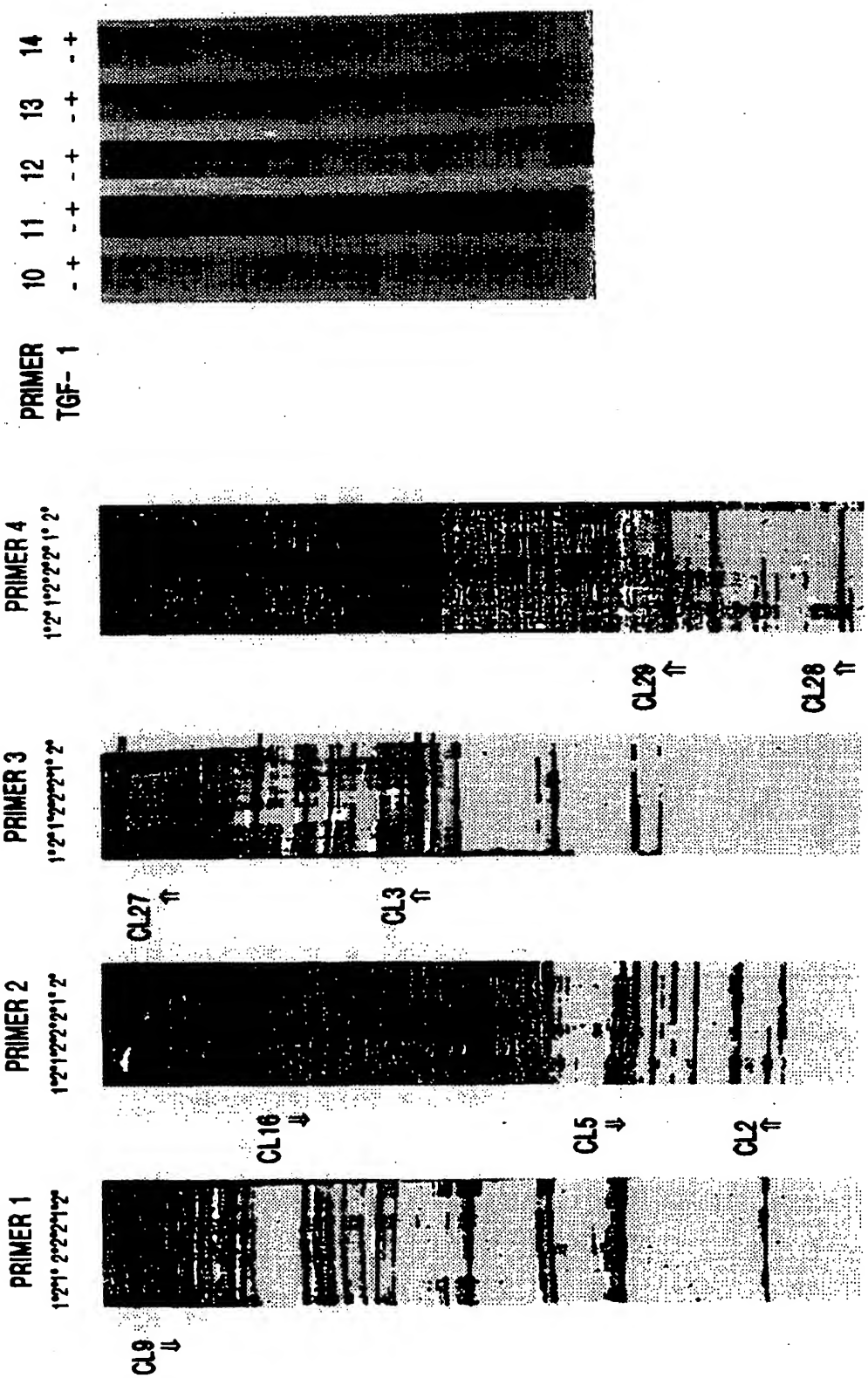
PRIMER TGF- 1
10 11 12 13 14
- + - + - + - + - +
PRIMER 4
1'2' 1'2'2'2' 1' 2'
CL28
CL29
PRIMER 3
1'2'1'2'2'1' 2'
CL3
CL27
PRIMER 2
1'2'1'2'2'1' 2'
CL2
CL5
CL16
PRIMER 1
1'2'1' 2'2'2'1'2'
CL9

FIG. 8

FIG. 9

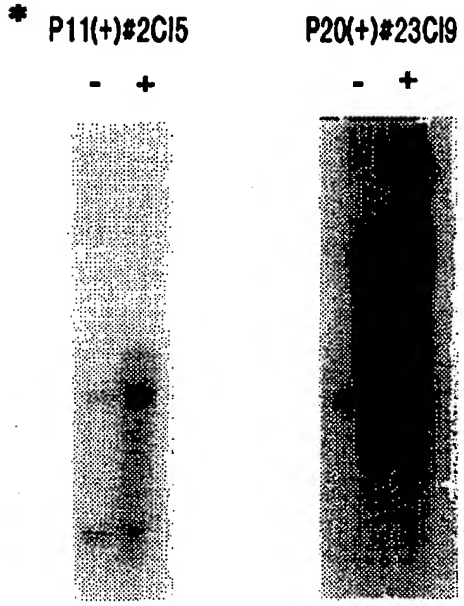

FIG. 10

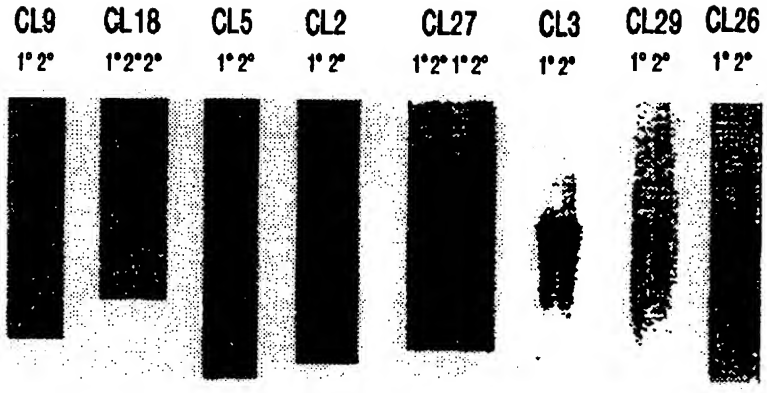

FIG. 11

CL-1#2

AATTTTTTTTTTCGACGGCCCAACGGAATTTTTTTTTTCGACGGCCCAACGGAATTTTT
TTTTTCGACGGCCCAACGGGAATTCGGCTTAGCTAAGGTCACCCAGACTTCATGGACT
TGTCTATTTTCTTGCCCAAAGGGATAGTTCCTCAGGTATTTGGGGACAGCATTCACCTC
TTGCAGGAGCTATGCCTGTGTGTTTGTGCTAAGTTGATACTTTCTGCGATGATCTCAC

(SEQ ID NO. 31)

CL-10#3

TACCATCGGAGAAAGAAGACCAAGCAAGGCTCAGGCAGCCACCGCCTGCTTCGCACT
GAGCCTCCTGACTCAGACTCAGAGTCCAGCACAGACGAAGAGGAATTTGGAGAATTG
GAAATCGCTCTCGTTTTGTCAAGGGAGACTATCCCGATGCTGCAAGATCTGCTGTCCCT
CTGGCCTTTGTCATCCTCGCGCCTGCGTTGTGGCCTCTGTGGGCTTGGTGTGGAGCAAA
TGGCTCTCAAGGAGGACTGAGTCTCAAGGAAATT (SEQ ID NO. 32)

CL-11A#5

AGCTAAGGTCAGGAGGTGTCTGAAGAATTGGCTGATGCATGGCAGGGATGTTGTTGAC
CTGCTTTTAGAACAATACTTCCATTTAATTATAGCATATCTTATGTGTGTATTAAAGCA
GAGCCGATCTGGTGGGGCTCATTAAGTAAATGTACTTACTGCAAAAGGTTCAACTGGT
GACCCCAGTTTTCCCCAGAAGCAATATGATAGGACAGAGGCGACTCCTGCAAGTTGTC
TCAGACTTCACACATACATTGTGACATTCTCTGAGCATGTGCACTGTACATGATATGAC
ACTATCAA (SEQ ID NO. 33)

CL-11C#2

AGCTAAGGTCCACTACCTTGTGAAGATGTATAAACACCTGAAATGTAGAAGCGATCCG
TATGTCAAGATCGAGGGGAAGGACGCTGACGACTGGCTGTGTGTGGACTTTGGGAGTA
TGGTGATCCATTTGATGCTTCCAGAAACCAGAGAAACCTATGAATTAGAGAAACTATG
GACTCTACGTTCTTTTGATGACCTTAGCTAAGCCGAATCAGCACACTGGCGGCGTTACT
AGTGGATCGAGCTCGTACAGCTGATGCATAGCTTGAGTATCTATAGGTTACTAATAGC
TGGCTATCATGTCAAGCGTTC (SEQ ID NO. 34)

FIG. 12A

CL-12#1

AGCTAAGGTCAAAATAAAAGCTCAAGATGACATCAGTCCCATTTGTCCTAAGTCCTGG
TGTTGTATGGATGGTAAGCAGCAGCCAATTATGGTGACAGGTGATAGATCCAATTTGT
TAACATTTCTCCATCTCTAAGCCATCCTTAAAGAAAATCATGAATGGAGTCACACCAT
CTTCACGGTAGTCCAGGAGAGCAACCATACCATCTGGATTCATGTTTCACCAATAAAA
ACTGGTAGTTATTGAATTAGCAAGGATGTGCTACTCTCTGCAGCTCAGC

(SEQ ID NO. 35)

CL-13#1

AGCTAAGGTCTCATGCAATGGAACTTAATTCTTAGAACTGTAAGAATTACATCAAACA
TAAAAGCCTCCCTATTAATGTAGTCCACAAAACTGGCAGGTATATATGCCTTCTGAAT
TTGTCTCCAGTGACTTTGGTAAATCTAACTAAATTTTTAAAAATTCTTAATGAATTTAT
CGTCAACAACAACCACCTCTTGGAAAATTAACCCTTGCAGTGTCTGTGTTAGACTCAG
AAGTCAA

(SEQ ID NO. 36)

CL-14#4

GAATTCGGCTTAGCTAAGGTCAGCGTGAAGTTTAAGCAGACATGAGTCTGAAACAGTC
TCATGACACATCTGATAGGATTTTTTAAGACTGCCTGGCTTAGTCTTACTGCTGTTAGT
GTATATTAGGTGTTGTACACATTATAAAGAAAATTATGTCTCATTATCTTGTTTAAGTC
AAGGAAAATAGAGAACTTTGGTCAAAT

(SEQ ID NO. 37)

CL-2#2

GAATTCGGCTTAGCTAAGGTCAGCGTGAAGTTTAAGCAGACATGAGTCTGAAACAGTC
TCATGACACATCTGATAGGATTTTTTAAGACTGCCTGGCTTAGTCTTACTGCTGTTAGT
GTATATTAGGTGTTGTACACATTATAAAGAAAATTATGTCTCATTATCTTGTTTAAGTC
AAGGAAAATAGAGAACTT

(SEQ ID NO. 38)

CL-2#3

GAATTCGGCTTAGCTAAGGTCAAAATACACGGATTGCAATCACTTTTCTAAACAAAAG
AAACAAAGTAACTGCTGAGGTTAGCAAAGATGAGTTCTCGTCATACTGCCTTGTACTG

FIG. 12B

TTTTGTGAACTGTGTTATTAAAAATCTGAGCTTAACAAAATCTTTACAAGTCACCTCAT
GAAAACAGCATTTGGCCAATAAGAGTTTAATTCCACACCAGTGAGACCTTAGCCT
(SEQ ID NO. 39)

CL-2#4

GAATTCGGCTTTCTGCGATCCACTCTTTGAAGCTATTGGCAAGATATTCAGCAACATCC
GCATCAGCACGCAGAAAGAGATATGAGGGACATTTCAAGGATGAAAGGTTTTTTTCCC
CCCTTACTATTTCCTTGGTGCCAATTCCAAGTTGCTCTCGCAGCAGCAAATTTATGAAT
GGTTTGTCTTGATCAAGAACAAAGAATTCATTCCCACCATTCTCATATATACTACTTTC
TCTTCTT (SEQ ID NO. 40)

CL-3#1

GAATTCGGCTTTCTGCGATCCACTCTTTGAAGCTATTGGCAAGATATTCAGCAACATCC
GCATCAGCACGCAGAAAGAGATATGAGGGACATTTCAAGGATGAAAGGTTTTTTTCCC
CCCTTACTATTTCCTTGGTGCCAATTCCAAGTTGCTCTCGCAGCAGCAAATTTATGAAT
GGTTTGTCTTGATCAAGAACAAAGAATTCATTCCACCATTCTCATATATCTACGTCTCT
TCTAG (SEQ ID NO. 41)

CL-4#1

GAATTCGGCTTTCTGCGATCCTAGAGCAGGTAAGTGAAGAAGGCCAGTAAGTTTTAAG
GATGGCCTTGTTGCCTTCTATCAAGTTCTCTGGGACTTTGTAATTTTGATTACTACTATT
GATACATGGTTATGGTCAGAAGGCCTCTTCTCCCTT (SEQ ID NO. 42)

CL-4#2

AGCTAAGGTCCGGACTCTATGGCATGACCCCAAAAACATTGGCTGGAAAGATTACACT
GCCTACAGGTGGCACCTGATTCACAGGCCTAAGACAGGCTACATGAGAGTCTTAGTGC
ATGAAGGAAAGCAAGTCATGGCTGACTCAGGACCAATTTATGACCAAACCTACGCTG
GTGGACGGCTGGGCTGTTTGTCTTCTCCAAGAGATGGTCTATTCTCGGACCTCAAGTAT
GAGTGCAGAGATGCTAGAGAGCAGGCTCAGTCTCAGCA
(SEQ ID NO. 43)

FIG. 12C

CL-5A#4

TGACCATCGAGTGCATCAGCCTCATCGGGCTGGCCGTCGGGAAGGAGAAATTCATGCA
GGATGCTTCAGATGTGATGCAGCTATTGTTGAAGACACAGACAGACTTCAATGATATG
GAAGATGACGACCCCCAGATTTCTTACATGATCTCAGCATGGGCCAGGATGTGCAAAA
TCTTGGGAAAGAATTCCAGCAGTACCTTCCCGTGGTTATGGGGCCGCTGATGAAGACT
GCTTCAATTAAGTCCTGAGTGCCTCTAGACACCAGGACATGAGATATGAGGTA

(SEQ ID NO. 44)

CL-6#2

TGACCATCGTGTAGTTGGTGTGCTTGTTGTCGAAGATGAGGGCCTCCTGGATGAGCTG
GTGCTGCTGCTCCAGCAGGTCCAGGCTGGGCTTGTAGTCCACGATGCTGCGCTCGTAC
TGCTTCAGGTGGCTCAGCTGGTCTTCCAGAGTCCCGTTCATCTCAATGGAGATGCGCCC
GATCTCCTCCATCTTAGTCTGGATCCACGGCCCCACCATATTGGCTTGGCTGGCGAACT
GTCGGCGAAGGCTGCATTGGATTGCT

(SEQ ID NO. 45)

CL-7#4

TGACCATCGAACACCCCAACACTCTCCACTACCTGCCATTTCTTCCAGCCTTATCCACA
CCACCCCGTTTCTCCTGAAGACTGATTTGCTTAGCAACTGCACTGAGCCAACCCTGAA
GACACATGATTATTGGTTGGGCTCCATTAAACAACAAGCCTAGTGCTTGGGAAGGGGG
GTGGGGAGGGGAAGAGACGTGAGAAGCATGTTGGCGTAGACCTTGAGGCATGGATGA
AGCATCTGCCGGCCTGACCTGGTACAGGTGGCATCTGCACTGCAGCAAGGC

(SEQ ID NO. 46)

FIG. 12D

CL-8#2

TGACCATCGAAGTGCAAAGGAAATGACTTGATTTCATGAAGTATCTCCAGAAGTAACG
CTTTGTTTTCTGCATCCTGAACTTTATTCCCAGTGAAGAGCTGAAAATCTGGACGCTCA
AAAAATGGAAGCACTTTGGAGAGAGCCCTTAACTCTATCAGGTACAGGAAGTACAAG
TTCCTCAGCCTTCGTGGGCCTTCTCCTTCAGTCAGAATCCATCAAAGGTGCTGGAACTC
TGTGACATTGTGACCCATTCTTTCAGCCAGTATCTGTAAGATAC

(SEQ ID NO. 47)

CL-9#1

GGGAACGAATGATCTGGAACTGTGGCTTGTAGACAACCCAAATATCTTAGGTAGGTAA
GAAATTCCAGCATCACACTATATAGGAAATACTGTGCGAAACTGACAGTTAACTGTGC
ACAAAGTTCAATGGCTTCAAAATAATGTATAAAGGATAAGAAGAAACCAGTTTACCAT
TTTGGT ATTATTTTGGTTGCTTTGTATAACTTCAATAATTT (SEQ ID NO. 48)

CL-54A#2.-SP

GGGAACGAATGATCTGGAACTGTGGCTTGTAGACAACCCAAATATCTTAGGTAGGTAA
GAAATTCCAGCATCACACTATATAGGAAATACTGTGCGAAACTGACAGTTAACTGTGC
ACAAAGTTCAATGGCTTCAAAATAATGTATAAAGGATAAGAAGAAACCAGTTTACCAT
TTTGGTATTATTTTGGTTGCTTTGTATAACTTCAATAATTT (SEQ ID NO. 49)

CL-54A#2.-S0

GACGTAAGCC (SEQ ID NO. 50)

CCACAAAGCAAGCTTCTGTCTGGAGTACAGCTCCTGTGACTATGGGTACCACAGGGCC
TTTGCGTGCACTGCACACACACAGGGATTGAGTCCTGGATGTTATGACACCTATGCGG
CAGACATAGACTGCCAGTGGATTGATATTACAGATGTACAACCTGGAAACTACATTCT
AAAGGTCAGTGTAAA

(SEQ ID NO. 51)

FIG. 12E

CTATCAATGAAGGGGGAGATCACTGGGTAAGTTCGAATGCCCTCAGGCAAGGTGGCC
CAGCCTTCCATTACTGAATTCAAAGATGGCACTGTTACTGTACGTTACTCACCCAGTGA
AGCTGGCCTGCATGAAATGGACATTCGCTATGACAATATGCATATCCCAGGAAGCCCT
CTGCAGTTCTATGTTGATTATGTCAACTGTGGCCACATCACTGCTTATGGTCC

(SEQ ID NO. 52)

TTAGCACCTCGACCACGAAATGAGGAAGATGCAACAGACGTGGTGGGCCTGGCTCAG
GCTGTAAACGCTCGGTCCCCACCTTCAGTAAAACAGAACAGCTTGGATGAAGACCTTA
TTCGGAAGCTAGCTTATGTTGCTGCTGGGGACCTGGCACCCATAAATGCTTTCATTGG
GGGCCTTGCTGCCCAGGAAGTCATGAAGGCCTGCTCTGGAAAGTTTATGCCCATCATG
CAGTGGTTGTACTTTGATGCTCTTGAATGTCTCCCAGAACGGACAAAGAGGCTCTGAC
AGAGGAGAGTGCCTCCCACGTCAGAACCGTTACGATGGGCAGGTAGCTGTATTGGTCA
GACTTCAGGAGAAGCTGAGAAGCAAA

(SEQ ID NO. 53)

TTAGCACCTCCAATGGCTGGGTACCAGCCAGCCGCAATGTCCGCTCCACAAATTTGGA
GTCTGTGAGGTACTGATTAACATTTTCTGCTGGCTGCTTGAAAAGGCCTTCAAATTCAT
CCCGGGCCCACTGAAGAGTGTGTTCGATGGCATTGGGAAAGTTTTTCAGGGTACAAAT
GGGGATGGATTTCTCTGGTGGATCCTGGCTAGACGTGATGGATTCTGTCAGGAAGGGG
ATTACCACCTGCACGTTGCCCTTT

(SEQ ID NO. 54)

TTAGCACCTCACACTCACATGCCCTTCTACATAGAGACTGGTTAAACAGCCCTCCCTCC
CTTGTCCCGACTTGACTTCCAGGCCCCTCTGCTTTCCTCTCACAACCACACCAGGTCTG
ATGGAGTCCAGTGCCTGCAGTGACCCAACATAGACTGCACTTTCACCTACCTACTGGA
TGGTCCTGCAGCCCAGACGGCTGCTCTTCTTTCTCATGGAGTTTCTCTCCTGCCTGAGA
TATGCTATCTGGTCTGCCCCTGTGTAGCTCCCATGGGATCCCTTAAAATCGATCCTTTT
TTAA

(SEQ ID NO. 55)

FIG. 12F

TTAGCACCTCGTGAGGAGACTGTTGTCCACAGGCCAGCTAGTGGTACCCTACTGAGAA
GTTGGGTTTTGGTTTTGTTTCCCTTGAAGGGTCGCTGTTAGAGGATGGAAGTAACTTCT
AATTCTTGATCTGTTTGTTGGTCTTGTTTTCAGTACTTTTTGCCAGTTGTATACACTTGG
AGAGGGAATTTGTATGCCTGTAATCTTGTTCTTGAGGTCAGAAATTCAAAACATTGGG
AGCTTTTGTTGTAAAGGTTAAACTGTGAATCCATATAGCAAATGCAGATCCTTTTACA
GTGTAAACCACATTTCCTGCCTCAGCCTAAAGCACTGGTCATTT (SEQ ID NO. 56)

ACCTGCATGCCTAAAGGAGTAGGCTTAGGGGTGGGGAGAGAGAAGGCATAGGCTTTT
CTAGTTATACAAAGCTGTGTAAGGCAAGGTTCCTTTCTACTAAATGGTCAGCTGTCACT
ACATTTATACTTTTGTATGTCATAAACCCTTTCTTTCATTCCTCCCTGGGTAACCAGGA
CAATCGGAGGGCAGTGTGTTACTGGGATTAGAGGACTAGCAATACTGGGTAACCCGCC
TAAGCTGGAAGGTGACGTAATACGTTTCTTTAAAGATTCAGTCAGTCAAGCAGTTTAG
CAATATCAAAATGTCTGGCTGTTTGGTCCAGTGTACACTGTT (SEQ ID NO. 57)

GCTATCTGCGAAACTACAGAAAGGAAGACAGCTTGGCCCAGCGCGGTGAAGTTCAGA
ATTCACTAGGTAGTTGTTGTTGGTTGACTTGGAGGTAGCTGGGTAATCAACAGCTTTCA
CTTTAGATTCAATGTGAACCGCAGAGTTACTCATGACCAAGAGTCTGGCAAACTCATT
AATGCTGTTTAATACTTGTTTGATATTTTTTCACCTTTTGAGCCCTTTTCCCAAAGAATT
CAATATCAGTTTAGTAGCAACAGTACAGTTGCCATTTAAATTGGTTTAGTTGCAGTATA
GCA (SEQ ID NO. 58)

GCTATCTGCGAAACTACAGAAAGGAAGACAGCTTGGCCCAGCGCGGTGAAGTTCAGA
ATTCACTAGGTAGTTGTTGTTGGTTGACTTGGAGGTAGCTGGGTAATCAACAGCTTTCA
CTTTAGATTCAATGTGAACCGCAGAGTTACTCATGACCAAGAGTCTGGCAAACTCATT
AATGCTGTTTAATACTTGTTTGATATTTTTTCACCTTTTGAGCCCTTTTCCCAAAGAATT

FIG. 12G

CAATATCAGTTTAGTAGCAACAGTACAGTTGCCATTTAAATTGGTTTAGTTGCAGTATA
GCA

(SEQ ID NO. 59)

GCTATACTGCAACTAAACCAATTTAAATGGCAACTGTACTGTTGCTACTAAACTGATA
TTGAATTCTTTGGGAAAAGGGCTCAAAAGGTGAAAAAATATCAAACAAGTATTAAAC
AGCATTAATGAGTTTGCCAGACTCTTGGTCATGAGTAACTCTGCGGTTCACATTGAATC
TAAAGTGAAAGCTGTTGATTACCCAGCTACCTCCAAGTCAACCAACAACAACTACCTA
GTGAATTCTGAACTTCACCGCGCTGGGCCAAGCTGTCTTCC

(SEQ ID NO. 60)

GCTATACTGCCCACCACATTGCCACACTCGGAATGACATTTCTATATTTTCACCTCCCC
AGATTTCCATTTCTTCATCGTAACTTCCAATGTGCTCAAAATATTTTTTAGATATAGAA
AAAAGGCCTCCTGCAAAGGTGGGGGTCTTAATTGGGTAGGTTTCATCTTTCCTTCTTTG
CTTCTCATGATCAGGAAGTGACTCCCAGCCAAAGGAAAGGCTCCAGTCAAAATTTCCA
CGGTTATGGTTGCTTCCGTACGGAGAAGGCTTGTTGAATTCAAATGTGTTTAGATCTAT
GGATGCGATGTCTGGACTCACCACGGCA

(SEQ ID NO. 61)

GCTATACTGCTGAAGGAGATCATTTTGGTGGATGATGCTAGTGTAGACGACTACCTGC
ATGAAAAGCTGGAGGAATACATAAAACAGTTTTCTATTGTGAAAATAGTCAGGCAGC
AAGAAAGGAAAGGCCTGATCACCGCGCGGTTGCTAGGGGCAGCTGTAGCAACTGCCG
AGACGCTCACGTTCTTAGATGCTCACTGTGAGTGCTTCTATGGCTGGCTGGAACCTCTG
CTGGCCAGGATAGCTGAGAACTACACTGCCG

(SEQ ID NO. 62)

AGTTGCCAGGGGGCAGCTCACGGCGCAGCTCATCCTCTGTGATGTAATTCTTATCTCC
AGCCAGGATCTTGAAGGAAGCCATGACCTGATCTGCAGTATCAGTATCTGCCGTCTCT

FIG. 12H

CGGGACATAAAGTCGATGAAGGCCTGGAACGTCACTACCCCCAAGCGGTTGGGGTCT
ACAATGCTCATGATTCGGGCAAACTCTGCCTCTCCCATGTTGTAACCCATGGAGATAA
GGCAGGCGCGGAAATCGTCTGTGTCCATCATGCCCGTCTTCTTCCGGTCAAAGTGGTT
GAAAGA (SEQ ID NO. 63)

AAGCCGTGTCGCTGAACTGGGAGGACACACTGCTCACCCTAGAAGGCTCTGGCTGACC
CTCCGCCCGGTTAAACAGGGACTTTGTGGCCATGTGCTGGCGACACAGGTCCTGGTAC
TCAAAAGTAGTGTCACCATGGGCCCCCTCCGGCCCCAGCGCTGCCAGGCGTCCTTATC
CCGCTGTCTCGAATGATGGCGCATACCAAGGCCACTGAAAGCCACTAGCAGCCCAGCG
ACGCCTGCCAGGGCCACTAGAGTAAGCAGCACTGAGCGCATGGGAGATATGCCAT
(SEQ ID NO. 64)

AAGCCGTGTCTGGACGTCCGTGTGTCCGGCTCTTGCTCACGCAGTCATGGCCTCCGGA
ACGCGCAAATCGGAAAGTCGGCTCCTGACTTCACGGCCACAGCGGTGGTGGATGGTGC
CTTCAAGGAAATCAAGCTTTCGGACTACAGAGGGAAGTACGTTGTCCTCTTTTTCTACC
CACTGGACTTCACTTTTGTTTGCCCCACGGAGATCATCGCTTTTAGCGACCATGCTGAG
GACTTCCGAAAGCTAGGCTGCGAGGTGCTGGGAGTGTCTGTGGACTCTCAGTTCACCC
ACCTGGCGTGGATCAATACCCCACGGAAAGAGGGAGGCTT (SEQ ID NO. 65)

AAGCCGTGTCGGAGGGCACCAAGGCTGTCACCAAGTACACCAGCTCCAAGTGAGTGC
TCAAGACTCAGCTCTTAACCCAAAGGCTCTTTTCAGAGCCACTCAAGACTTCAAAATT
GGAGCTTTAATGCTGACTTAGTGACTACCGGGAAAATAACTGACTTCATCTGCAGGAT
TGTGTACAAACACTTATGGTTTAGTAAATCGAAAAGATAGACATTGCCCATCAGTTCT
GTCTGGTCCACTTAAATATGCTTTTTTCTTAGAAGTTCTAAGAACCCTGTCAATAACCT
ATCTAGGTCCAGTCCTTGAGTTCAAAGGCCAAATACCAATG (SEQ ID NO. 66)

FIG. 12I

CAACGCTCAGGATGTAAGCTGTTTCCAGCACCTGGTTCAAGCGAATGTAAGAAATAAG
AAGGTGTTGAAAGATGCCGTGAATAACATTACAGCAAAGGGGATCACAGATTACAAG
AAAGGCTTTAGCTTTGCCTTCGAACAGCTACTTAATTATAATGTTTCCAGAGCTAATTG
CAATAAGATTATCATGTTATTCACGGATGGAGGAGAAGAGAGAGCCCAGGAGATATT
TGCCAAATACAATAAAGACAAAAAAAGTCCGTGTGTTTACATTTTCCGTCGGTCAACAT
AATTATGACAGAGGACCTATTCAGTGGATGGCTTGTGAAATAAAGGTTACTATTATGA
GATTCCTCCATT

(SEQ ID NO. 67)

TCAACGCTCATCACACCAAGAATCAACTGGTTCTTCAAGTTTGTCTTATTTTCAGATTG
GCCAGTGACGTTGAAGACTGGTAGAGTTCCAGTAATGACAAGTCCCAGTTCCAGGGCA
TCCAAATACACATTTGTCCATTGAACTTGCTTCGCTTTGTCACCAGCTAAAACCATTGG
TCTTCCCAGAACATCTAGATATTCCTGAGTATTGATTCTTATTGCACCAATGGAGGGAA
TCTCATAATAGTAACCTTTATTTTCACAAGCCATCCACTGAATAGGTCTCTGTCATAAT
TATGTTGACCGACGGAAATGTAA

(SEQ ID NO. 68)

TAACGCTCAGGAGAAGAATAGGAATGCAGAGAACTCTGCCACAGCCCCCACGCTCCC
GGGCAGCACCTCAGCCACCACCGCAACCACCACCCCTGCTGTAGATGAAAGCAAGCCT
TGGAACCAGTATCGCTTGCCTAAGACTCTTATACCTGACTCCTACCGGGTGATCTTGAG
ACCCTACCTCACCCCCAACAATCAGGGCCTGTACATCTTCCAAGGCAACAGTACTGTT
CGCTTTACCTGCAACCAGACCACGGATGTCATTATCATCCACAGCAAAAAGCTCAACT
ACACCCTCAAAGGAAACCACAGGGTGG

(SEQ ID NO. 69)

CGAGTCAGACGGCTTCAGCATCGAGACCTGTAAGATCATGGTGGACATGCTGGATGAA
GATGGGAGTGGCAAGCTTGGCCTGAAGGAGTTCTACATCCTCTGGACGAAGATTCAGA
AATACCAAAAAATCTACCGGGAAATCGATGTGGACAGGTCTGGAACTATGAATTCCTA

FIG. 12J

CGAGATGCGGAAAGCACTGGAAGAAGCAGGTTTCAAGCTGCCCTGTCAACTCCATCA
AGTCATCGTTGCCCGGTTTGCAGACGACGAGCTAATCATCGACTTTGACAATTTTG
(SEQ ID NO. 70)

CGAGTCAGACAACCTGTTCAAGTGGGGTGGGGACCATCCACGGAGCAGCCGGCACCG
TATATGAAGACCTGAGGTACAAACTCTCCCTAGAGTTCCCCAGCGGCTACCCTTACAA
CGCACCCACAGTGAAGTTCCTCACACCCTGCTACCACCCCAACGTGGACACCCAGGGC
AACATCTGCCTGGACATCCTCAAGGATAAGTGGTCTGCACTATATGATGTCAGGACTA
TCTTGCTCTCTATCCAGAGCCTGCTAGGAGAACCCAACATCGATAGCCTTTGAACACA
CACGCTGCGGAACTCTGGAAAA (SEQ ID NO. 71)

TATGAGTCCGGAGCGACGGCTACGAGTGTGAACTGTTCCAGCCCCGAGCGACACACCA
GAAGTTATGACTACATGGAAGGAGGGGATATAAGGGTGAGAAGACTGTTCTGTCGCA
CCCAGTGGTACCTGAGGATTGACAAACGAGGCAAAGTGAAAGGGACCCAGGAGATGA
AGAACAGCTACAACATCATGGAAATCAGGACCGTGGCAGTTGGAATTGTGGCAATCA
AAGGGGTGGAAAGTGAATACTATCTTGCCATGAACAAGGAAGGGAAACTCTATGCAA
AGAAAGAATGCAATGAGGATTGCAACTTCAAAGAACTGATTCTGGAAAACCATTATA
ACACCTATG (SEQ ID NO. 72)

TATGAGTCCGAGGAGGAGCACAATGCTGGGAGTGTGGAAAGCCAGGTTGTCCCCAGC
ACACACCGAGTGACCGATTCCAAGTTCCATCCACTCCATGCCAAGATGGATGTCATCA
AAAAAGGCCACGCCAGGGACAGCCAGCGCTACAAAGTTGACTATGAGTCTCAAAGCA
CAGACACCCAGAACTTCTCCTCCGAGTCTAAGCGGGAGACAGAATACGGTCCCTGCCG
CAGAGAAATGGAGGACACACTGAATCATCTGAAGTTCCTCAATGTGCTGAGTCCAGAG
TCTCACATCCAAACTGTGACAAGAAGGGG
(SEQ ID NO. 73)

FIG. 12K

TCGCCCGGGACTTCATGCGATTGAGAAGATTGTCTACCAAATATAGAACAGAAAAGAT
TTATCCCACAGCCACTGGAGAAAAAGAAGAAAATGTTAAAAAGAACAGATATAAGGA
CATACTGCCATTTGATCACAGCCGAGTTAAGTTGACTTTGAAGACTCCATCCCAAGAT
TCAGATTATATCAATGCAAATTTTATTAAGGGTGTGTATGGGCCAAAAGCATATGTGG
CAACCCAAGGGCCTTT

(SEQ ID NO. 74)

TGTGGAAAGCCAGGTTGTCCCCAGCACACACCGAGTGACCGATTCCAAGTTCCATCCA
CTCCATGCCAAGATGGATGTCATCAAAAAAGGCCACGCCAGGGACAGCCAGCGCTAC
AAAGTTGACTATGAGTCTCAAAGCACAGACACCCAGAACTTCTCCTCCGAGTCTAAGC
GGGAGACAGAATACGGTCCCTGCCGCAGAGAAATGGAGGACACACTGAATCATCTGA
AGTTCCTCAATGTGCTGAGTCCAGAG

(SEQ ID NO. 75)

TGACCATCGAAGTGCAAAGGAAATGACTTGATTTCATGAAGTATCTCCAGAAGTAACG
CTTTGTTTTCTGCATCCTGAACTTTATTCCCAGTGAAGAGCTGAAAATCTGGACGCTCA
AAAAATGGAAGCACTTTGGAGAGAGCCCTTAACTCTATCAGGTACAGGAAGTACAAG
TTCCTCAGCCTTCGTGGGCCTTCTCCTTCAGTCAGAATCCCATCAAAGCGCTGCTGGAA
CTCTGTGACATTGTGACCCCATTTCTTTTCCAGCCAAGTATCTTGTAAAAGATACCTTG
CACTCAAATGCACATTAATGCTTGCGTGCAGGCCAGATATAAGTCTGTAGAATCGCTC
TTTCTACACAGAGGCCTTCTAGCCAGTTGTAAA

(SEQ ID NO. 76)

CTGCTTGATGCTAAGCCCGGCAGCCTGTGTTTCATCTACAGGATGCACAACATAAAAG
AAAAGATCTGATTCCCGCAGGTTCTCTTCTGACCTACACACACACACACTAAAATAAC
ATTTAAAAATATGTGCCAAATTATATTTGTTCGGGTGCCACCTTCCACCAGCTTACCAC
TACGGTAGAACTGTCAAATTCATCTCCCTGAATTTGTCTTAAAGGGGTGTCCATGCAC
AGGCCCAAGAGTCACCTCCAATGAAATAAATGTAATACTGAAGTATGCCATGATGTTT

FIG. 12L

GTTGTTTTCTTTCATCGTAAGCCTGTAAGCAGGAAAAATAGTAATAGATAGAATAGAG
ACTTACCAGTGGTCGATGGCCTGGTCAGTCTGTGCGGTGACTAGGACCAGG

(SEQ ID NO. 77)

ACCTGCATGCCGAGTGTGACGCCTTTGAGGAGAAGATCCAGGCTGCCGGAGGGATCG
AACTCTTTGTCGGAGGCATTGGCCCCGATGGACACATTGCCTTCAATGAGCCAGGCTC
CAGCCTGGTGTCCAGGACCCGTGTGAAGACTCTGGTTATGGACACCATCCTGGCCAAC
GCTAGGTTCTTTGATGGTGATCTTGCCAAGGTGCCCACCATGGCCCTGACAGTGGGTG
TCGGCACTGTCATGGATGCTAAAGAGGTGATGATCCTCATCACAGGCGCTCACAAGGC
CTTTGCTCTGTACAAAGCCATCGATGGAGGCGTGAACCACATGTGGACGGTGTG

(SEQ ID NO. 78)

GCTATACTGCAATGTTAGGGGAATGAACGCGTTTTCCTACTGCACTGGGGACTTTTAG
ATAGGTTAATGAAAGGCCTTTTATTCTGTTACTGGACACGAAAACTTTGTCTAATTTCT
TATACTCTATTGTACGTTTACAGTCGCAGCACTAAAATGGAAGACATCAAACATTTTT
AACAGAAAAAAAAAAAAGATGTAAAAACTAACTAAGGACTATTTATTGATAATGTTTTG
CTACTCCTGTCAGACAATGGCTATAAACTGAATTAGGCAGTCTTAAAAAAAAAAAAAG
AAAAAAAAGAAAAAAGAAAAAAAGAAAAGAAAAGAAAAAAAACTGG

(SEQ ID NO. 79)

AGCTAAGGTCGGGTACTCTGATACTTCAGAGTTTAAAATCATCAGCCCTTGTAGATCT
ATTCCTAAATCTTATGAAAATGCTCAGATGTTTACACAGCTGTGAAACAGGGTCAGTT
CAGATCGCTGATGGCTTGAGAATGTGTTTCTTGTTGACATCAGGAACTGGAAATGTTT
ACTTCCCGTCATTTATGAGTCATCAAGTATCTCGGCTCTTTTAAGAGCGCAAGATAAA
ACAAGCTTAAACCAGGTGATAAGAGCAGAGTCCACTTGAGTCTGAGCTCACCCGAGA
ACTTGCTATCGAGGACATTTGGAATGGGAGTGTGCAGGCTTCCTTCAGTTACTGAATG
AGTCCATCTGCTAGTCACCTTGAC

(SEQ ID NO. 80)

FIG. 12M

AGCTAAGGTCCAGGGGGCAAAGCGGTGACGTGTGCACATCGATATGAGAAACGGCAG
CACGTCAACACGAAGCAGGAGTCGCGGGATATCTTTGGAAGATGTTATGTCCTAAGTC
AGAATCTCAGAATTGAAGATGATATGGACGGAGGAGACTGGAGTTTCTGCGATGGCC
GGTTGAGAGGCCATGAAAAGTTTGGCTCCTGTCAGCAAGGAGTAGCGGCTACTTTCAC
TAAGGACTTTCATTACATTGTTTTTGGAGCCCCAGGGACTTACAACTGGAAAGGGATC
GTCGTGTAGAACAAAAGAATAACACTTTTTT

(SEQ ID NO. 81)

AAGCCGTGTCTGTGCTCAAGGAAGAAACCCACTGGACCAACTTCTGTCAGAAAGGAA
AACCTTGTTCAAAGTTTCAGGACCCTGTTCTTTGCTTATTTGCACATGGTCACCTTGGT
CTGAGCTAGCCACCATTGTCACCCACAGCTGCAAAGAAAGCAGACCTTAGGAAACACT
GTCACGGCTGAGTGTGACTGCCTTGTTCATCCCCTGGACTGGTACTGTGTTGCCTGCAG
TACCATTGGGATCCCATAGCAAGAGAGGGAGAGGGAGATGTTAGTTAGCCTTTGCTAC
GAACCAAGCTGTCCCAAGTCTCAACAGCTAAACAGGTATTCATTTACCATGATTCTAT
GGTTAGCTAAGCTCTTGAG

(SEQ ID NO. 82)

CTTTCTACCCTGGAGGATGTGCTTGAGGCACACTGCTCCTGTGCTCTCCACTTGAGGCA
TAAGCCCAGTCAGTTGTGCATAGATGATTAACCTCTGACCCCTAAAGATGGTAAGTTG
CTCTGGAGAAAGCATTTTAACAGACAAACCAGGAGGCAAATCCCAACTTAGAGAGAT
GTTATCCACTGCACACTGTAGAGCAAACTTGAGAGACCCAAGAGCCTTGGTCTGCATC
CTGTCCTTGCCTGTGATAAACACTCGAGTACCCCCTGATACCGGGCGATATTTTTGATT
AACTGGTCGAGGCTCCTTGTCCAATTCCAAAAGAGAACATCTGTGTTTC

(SEQ ID NO. 83)

FIG. 12N

TGGTAAAGGGCATCTGTAAATACACTCTATGAGGAAATTAAAACTTGAACATGGCAGT
CTGACATTGCAAAACAAAACAAAACAAAACTGACCCTCCAATAGCAGCGAAAACAAC
GTGAAAGATACAAAGCAATGAGAATCTGGTTCTGAACGCCTGGGATCCTGGGAGTCAT
CGGTAGCAGCGCCATGAGAGGAGCCGTGGCCTGTCCCATGTGGTCCCACCTTCACCTC
TTCCCTCACATCCCTCTTAAG

(SEQ ID NO. 84)

TGGTAAAGGGGGCAAGGGCAAAGGCACGGGAGACAGAGGCCACTGCATCTGTACCCA
CATCAGACATGTTTGTCCATTTTCTCTCATTTGGCCTTAGACCATTGGCAAGAGTAAAT
GCTCTTAGTCCCGTTATCTAGAAATTTCTTCCTTTGGGGAGAACCACTTATAGACAATA
TCAGCTCTCTACAAATAACACGAAAGGTCGTAACAC
AGCAAGTGACCAGAAAGTGCCCGTCCTTGCGGCTCTGATCCACGTGGCTCTCCGTAGA
CAAATTGTTTTTTCTTGTAGGGATATCTGTTTTGCTTCTGAACTTTCTTACAAGTGTTTG
GGACTCTTCGGGTGGCGTT

(SEQ ID NO. 85)

TGGTAAAGGGTCAAGTGTTCGATCAGAGTGGAGCTCCATTACCGAATGTAATCGTGGA
AGTCCAAGACAGAAAGCATATCTGCCCGTTTAGAACCAACAAGCTTGGAGAATACTAT
CTGCTTCTGCTGCCCGGGTCCTACGTGATCAATGTTACAGTCCCTGGACACGACTCCTA
CCTCACGAAGCTTACTATTCCAGGGAAATCCCAGCCCTTCAGTGCTCTTAAAAAGGAT
TTTCACCTCCCGCTGCGATGGCAGCCGGATTCCATCTCCGTATCCAATCCTTCGTGCCG
ATGATTCCGCTGTACAAATTCATGCCAAGCCACTCGGCTGCCACAAAGCCTAGTCTGG
G

(SEQ ID NO. 86)

GAATTCGGCTTTCTGCGATCCACTCTTTGAAGCTATTGGCAAGATATTCAGCAACATCC
GCATCAGCACGCAGAAAGAGATATGAGGGACATTTCAAGGATGAAAGGTTTTTTTCCC
CCCTTACTATTTCCTTGGTGCCAATTCCAAGTTGCTCTCGCAGCAGCAAATTTATGAAT

FIG. 12O

GGTTTGTCTTGATCAAGAACAAAGAATTCATTCCCACCATTCTCATATATACTACTTTC
TCTTCTT

(SEQ ID NO. 87)

GAATTCGGCTTTCTGCGATCCACTCTTTGAAGCTATTGGCAAGATATTCAGCAACATCC
GCATCAGCACGCAGAAAGAGATATGAGGGACATTTCAAGGATGAAAGGTTTTTTTCCC
CCCTTACTATTTCCTTGGTGCCAATTCCAAGTTGCTCTCGCAGCAGCAAATTTATGAAT
GGTTTGTCTTGATCAAGAACAAAGAATTCATTCCACCATTCTCATATATCTACGTCTCT
TCTAG

(SEQ ID NO. 88)

ACGAGGGGAAACCTCCTCAGAGCCTGCAGCCAGCCACGCGCCAGCATGTCTGGGGGC
AAATACGTAGACTCCGAGGGACATCTCTACACTGTTCCCATCCGGGAACAGGGCAACA
TCTACAAGCCCAACAACAAGGCCATGGCAGACGAGGTGACTGAGAAGCAAGTGTATG
ACGCGCACACCAAGGAGATTGACCTGGTCAACCGCGACCCCAAGCATCTCAACGACG
ACGTGGTCAAGATTGACTTTGAAGATGTGATTGCAGAACCAGAAGGGACACACAGTTT
CGACGGCATCTGGAAGGCCAGCTTCACCACCTTCACTGTGACAAAATATTGGTTTTAC
CGCTTGTTGTCTACGATCTTCGGCATCCCAATGGCACTCATCTGGGGCATTTACTTTGC
CATTCTCTCCTTCCTGCACATCTGGGCGGTTGTACCGTGCATCAAGAGCTTCCTGATTG
AGATTCAGTGCATCAGCCGCGTCTACTCCATCTACGTCCATACCTTCTGCGATCCACTC
TTTGAAGCTATTGGCAAGATATTCAGCAACATCCGCATCAGCACGCAGAAAGAGATAT
GAGGGACATTTCAAGGATGAAAGGTTTTTTTCCCCCCTTACTATTTCCTTGGTGCCAAT
TCCAAGTTGCTCTCGCAGCAGCAAATTTATGAATGGTTTGTCTTGATC

(SEQ ID NO. 89)

MECLYYFLGFLLLAARLPLDAAKRFHDVLGNERPSAYMREHNQLNGWSSDENDWNEKL
YPVWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIF
PRCQKEDANGNIVYEKNCRNEAGLSADPYVYNWTAWSEDSDGENGTGQSHHNVFPDGK

FIG. 12P

PFPHHPGWRRWNFIYVFHTLGQYFQKLGRCSVRVSVNTANVTLGPQLMEVTVYRRHGRA
YVPIAQVKDVYVVTDQIPVFVTMFQKNDRNSSDETFLKDLPIMFDVLIHDPSHFLNYSTIN
YKWSFGDNTGLFVSTNHTVNHTYVLNGTFSLNLTVKAAAPGPCPPPPPPPRPSKPTPSLGP
AGDNPLELSRIPDENCQINRYGHFQATITIVEGILEVNIIQMTDVLMPVPWPESSLIDFVVTC
QGSIPTEVCTIISDPTCEITQNTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLAL
TSTLISVPDRDPASPLRMANSALISVGCLAIFVTVISLLVYKKHKEYNPIENSPGNVVRSKGL
SVFLNRAKAVFFPGNQEKDPLLKNQEFKGVS

(SEQ ID NO. 90)

1 CAGATGCCAG AAGAACACTG TTGCTCTTGG TGGACGGGCC CAGAGGAATT CAGAGTTAAA

61 CCTTGAGTGC CTGCGTCCGT GAGAATTCAG CATGGAATGT CTCTACTATT TCCTGGGATT

121 TCTGCTCCTG GCTGCAAGAT TGCCACTTGA TGCCGCCAAA CGATTTCATG ATGTGCTGGG

181 CAATGAAAGA CCTTCTGCTT ACATGAGGGA GCACAATCAA TTAAATGGCT GGTCTTCTGA

241 TGAAAATGAC TGGAATGAAA AACTCTACCC AGTGTGGAAG CGGGGAGACA TGAGGTGGAA

301 AAACTCCTGG AAGGGAGGCC GTGTGCAGGC GGTCCTGACC AGTGACTCAC CAGCCCTCGT

361 GGGCTCAAAT ATAACATTTG CGGTGAACCT GATATTCCCT AGATGCCAAA AGGAAGATGC

421 CAATGGCAAC ATAGTCTATG AGAAGAACTG CAGAAATGAG GCTGGTTTAT CTGCTGATCC

481 ATATGTTTAC AACTGGACAG CATGGTCAGA GGACAGTGAC GGGGAAAATG GCACCGGCCA

541 AAGCCATCAT AACGTCTTCC CTGATGGGAA ACCTTTTCCT CACCACCCCG GATGGAGAAG

FIG. 12Q

601 ATGGAATTTC ATCTACGTCT TCCACACACT TGGTCAGTAT TTCCAGAAAT TGGGACGATG

661 TTCAGTGAGA GTTTCTGTGA ACACAGCCAA TGTGACACTT GGGCCTCAAC TCATGGAAGT

721 GACTGTCTAC AGAAGACATG GACGGGCATA TGTTCCCATC GCACAAGTGA AAGATGTGTA

781 CGTGGTAACA GATCAGATTC CTGTGTTTGT GACTATGTTC CAGAAGAACG ATCGAAATTC

841 ATCCGACGAA ACCTTCCTCA AAGATCTCCC CATTATGTTT GATGTCCTGA TTCATGATCC

901 TAGCCACTTC CTCAATTATT CTACCATTAA CTACAAGTGG AGCTTCGGGG ATAATACTGG

961 CCTGTTTGTT TCCACCAATC ATACTGTGAA TCACACGTAT GTGCTCAATG GAACCTTCAG

1021 CCTTAACCTC ACTGTGAAAG CTGCAGCACC AGGACCTTGT CCGCCACCGC CACCACCACC

1081 CAGACCTTCA AAACCCACCC CTTCTTTAGG ACCTGCTGGT GACAACCCCC TGGAGCTGAG

1141 TAGGATTCCT GATGAAAACT GCCAGATTAA CAGATATGGC CACTTTCAAG CCACCATCAC

1201 AATTGTAGAG GGAATCTTAG AGGTTAACAT CATCCAGATG ACAGACGTCC TGATGCCGGT

1261 GCCATGGCCT GAAAGCTCCC TAATAGACTT TGTCGTGACC TGCCAAGGGA GCATTCCCAC

1321 GGAGGTCTGT ACCATCATTT CTGACCCCAC CTGCGAGATC ACCCAGAACA CAGTCTGCAG

1381 CCCTGTGGAT GTGGATGAGA TGTGTCTGCT GACTGTGAGA CGAACCTTCA ATGGGTCTGG

1441 GACGTACTGT GTGAACCTCA CCCTGGGGGA TGACACAAGC CTGGCTCTCA CGAGCACCCT

FIG. 12R

1501 GATTTCTGTT CCTGACAGAG ACCCAGCCTC GCCTTTAAGG ATGGCAAACA GTGCCCTGAT

1561 CTCCGTTGGC TGCTTGGCCA TATTTGTCAC TGTGATCTCC CTCTTGGTGT ACAAAAAACA

1621 CAAGGAATAC AACCCAATAG AAAATAGTCC TGGGAATGTG GTCAGAAGCA AAGGCCTGAG

1681 TGTCTTTCTC AACCGTGCAA AAGCCGTGTT CTTCCCGGGA AACCAGGAAA AGGATCCGCT

1741 ACTCAAAAAC CAAGAATTTA AAGGAGTTTC TTAAATTTCG ACCTTGTTTC TGAAGCTCAC

1801 TTTTCAGTGC CATTGATGTG AGATGTGCTG GAGTGGCTAT TAACCTTTTT TTCCTAAAGA

1861 TTATTGTTAA ATAGATATTG TGGTTTGGGG AAGTTGAATT TTTTATAGGT TAAATGTCAT

1921 TTTAGAGATG GGGAGAGGGA TTATACTGCA GGCAGCTTCA GCCATGTTGT GAAACTGATA

1981 AAAGCAACTT AGCAAGGCTT CTTTTCATTA TTTTTTATGT TTCACTTATA AAGTCTTAGG

2041 TAACTAGTAG GATAGAAACA CTGTGTCCCG AGAGTAAGGA GAGAAGCTAC TATTGATTAG

2101 AGCCTAACCC AGGTTAACTG CAAGAAGAGG CGGGATACTT TCAGCTTTCC ATGTAACTGT

2161 ATGCATAAAG CCAATGTAGT CCAGTTTCTA AGATCATGTT CCAAGCTAAC TGAATCCCAC

2221 TTCAATACAC ACTCATGAAC TCCTGATGGA ACAATAACAG GCCCAAGCCT GTGGTATGAT

2281 GTGCACACTT GCTAGACTCA GAAAAAATAC TACTCTCATA AATGGGTGGG AGTATTTTGG

2341 TGACAACCTA CTTTGCTTGG CTGAGTGAAG GAATGATATT CATATATTCA TTTATTCCAT

FIG. 12S

2401 GGACATTTAG TTAGTGCTTT TTATATACCA GGCATGATGC TGAGTGACAC
TCTTGTGTAT

2461 ATTTCCAAAT TTTTGTATAG TCGCTGCACA TATTTGAAAT CATATATTAA
GACTTTCCAA

2521 AGATGAGGTC CCTGGTTTTT CATGGCAACT TGATCAGTAA GGATTTCACC
TCTGTTTGTA

2581 ACTAAAACCA TCTACTATAT GTTAGACATG ACATTCTTTT TCTCTCCTTC
CTGAAAAATA

2641 AAGTGTGGGA AGAGACAAAA AAAAAAAAA //

(SEQ ID NO. 91)

AAGGTGAAAGATGTGTATGTGATAACAGATCAGATCCCTGTATTCGTGACCATGTCCC
AGAAGAATGACAGGAACTTGTCTGATGAGATCTTCCTCAGAGACCTCCCCATCGTCTT
CGATGTCCTCATTCATGATCCCAGCCACTTCCTCAACGACTCTGCCATTTCCTACAAGT
GGAACTTTGGGGACAACACTGGCCTGTTTGTCTCCAACAATCACACTTTGAATCACAC
TTATGTGCTCAATGGAACCTTCAACCTTAACCTCACCGTGCAAACTGCAGTGCCCGGG
CCATGCCCTCCCCCTTCGCCTTCGACTCCGCCTCCACCTTCGTA

(SEQ ID NO. 92)

AAGGTGAAAGATGTGTATGTGATAACAGATCAGATCCCTGTATTCGTGACCATGTCCC
AGAAGAATGACAGGAACTTGTCTGATGAGATCTTCCTCAGAGACCTCCCCATCGTCTT
CGATGTCCTCATTCATGATCCCAGCCACTTCCTCAACGACTCTGCCATTTCCTACAAGT
GGAACTTTGGGGACAACACTGGCCTGTTTGTCTCCAACAATCACACTTTGAATCACAC
TTATGTGCTCAATGGAACCTTCAACCTTA

(SEQ ID NO. 93)

AAGGTGAAAGATGTGTATGTGATAACAGATCAGATCCCTGTATTCGTGACCATGTCCC
AGAAGAATGACAGGAACTTGTCTGATGAGATCTTCCTCAGAGACCTCCCCATCGTCTT

FIG. 12T

CGATGTCCTCATTCATGATCCCAGCCACTTCCTCAACGACTCTGCCATTTCCTACAAGT
GGAACTTTGGGGACAACACTGGCCTGTTTGTCTCCAACAATCACACTTTGAATCACAC
TTATGTGCTCAATGGAACCTTCAACCTTAACCTCACCGTGCAAACTGCAGTGCCCGGG
CCATGCCCTCCCCCTTCGCCTTCGACTCCGCCTCCACCTTCGTA (SEQ ID NO. 94)

TACGAAGGTGGAGGCGGAGTCGAAGGCGAAGGGGGAGGGCATGGCCCGGGCACTGCA
GTTTGCACGGTGAGGTTAAGGTTGAAGGTTCCATTGAGCACATAAGTGTGATTCAAAG
TGTGATTGTTGGAGACAAACAGGCCAGTGTTGTCCCCAAAGTTCCACTTGTAGGAAAT
GGCAGAGTCGTTGAGGA

(SEQ ID NO. 95)

AAGGTGAAAGATGTGTATGTGATAACAGATCAGATCCCTGTATTCGTGACCATGTCCC
AGAAGAATGACAGGAACTTGTCTGATGAGATCTTCCTCAGAGACCTCCCCATCGTCTT
CGATGTCCTCATTCATGATCCCAGCCACTTCCTCAACGACTCTGCCATTTCCTACAAGT
GGAACTTTGGGGACAACACTGGCCTGTTTGTCTCCAACAATCACACTTTGAATCACAC
TTATGTGCTCAATGGAACCTTCAACCTTAACCTCACCGTGCAAACTGCAGTGCCCGGG
CCATGCCCTCCCCCTTCGCCTTCGACTCCGCCTCCACCTTCGTA

(SEQ ID NO. 96)

RRWRRSRRRRGRAWPGHCSLHGEVKVEGSIEHISVIQSVIVGDKQASVVPKVPLVGNGRV
VEEVAGIMNEDIEDDGEVSEEDLIRQVPVILLGHGHEYRDLICYHIHIFHL

(SEQ ID NO. 97)

KVKDVYVITDQIPVFVTMSQKNDRNLSDEIFLRDLPIVFDVLIHDPSHFLNDSAISYKWNFG
DNTGLFVSNNHTLNHTYVLNGTFNLNLTVQTAVPGPCPPPSPSTPPPPS (SEQ ID NO. 98)

FIG. 12U

YEGGGGVEGEGGGHGPGTAVCTVRLRLKVPLST*V*FKV*LLETNRPVLSPKFHL*EMAES
LRKWLGS*MRTSKTMGRSLRKISSDKFLSFFWDMVTNTGI*SVITYTSFT (SEQ ID NO. 99)

MECLYYFLGFLLLAARLPLDAAKRFHDVLGNERPSAYMREHNQLNGWSSDENDWNEKL
YPVWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCQKEDANGNIVYE
KNCRNEAGLSADPYVYNWTAWSEDSDGENGTGQSHHNVFPDGK
PFPHHPGWRRWNFIYVFHTLGQYFQKLGRCSVRVSVNTANVTLGPQLMEVTVYRRHGRA
YVPIAQVKDVYVVTDQIPVFVTMFQKNDRNSSDETFLKDLPIMFDVLIHDPSHFLNYSTIN
YKWSFGDNTGLFVSTNHTVNHTYVLNGTFSLNLTVKAAAPGPCPPPPPPPRPSKPTPSLGP
AGDNPLELSRIPDENCQINRYGHFQATITIVEGILEVNIIQMTDVLMPVPWPESSLIDFVVTC
QGSIPTEVCTIISDPTCEITQNTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLAL
TSTLISVPDRDPASPLRMANSALISVGCLAIFVTVISLLVYKKHKEYNPIENSPGNVVRSKGL
SVFLNRAKAVFFPGNQEKDPLLKNQEFKGVS (SEQ ID NO. 100)

1 CAGATGCCAG AAGAACACTG TTGCTCTTGG TGGACGGGCC CAGAGGAATT
CAGAGTTAAA

61 CCTTGAGTGC CTGCGTCCGT GAGAATTCAG CATGGAATGT CTCTACTATT
TCCTGGGATT

121 TCTGCTCCTG GCTGCAAGAT TGCCACTTGA TGCCGCCAAA CGATTTCATG
ATGTGCTGGG

181 CAATGAAAGA CCTTCTGCTT ACATGAGGGA GCACAATCAA TTAAATGGCT
GGTCTTCTGA

241 TGAAAATGAC TGGAATGAAA AACTCTACCC AGTGTGGAAG CGGGGAGACA
TGAGGTGGAA

301 AAACTCCTGG AAGGGAGGCC GTGTGCAGGC GGTCCTGACC AGTGACTCAC
CAGCCCTCGT

361 GGGCTCAAAT ATAACATTTG CGGTGAACCT GATATTCCCT AGATGCCAAA
AGGAAGATGC

FIG. 12V

421 CAATGGCAAC ATAGTCTATG AGAAGAACTG CAGAAATGAG GCTGGTTTAT CTGCTGATCC

481 ATATGTTTAC AACTGGACAG CATGGTCAGA GGACAGTGAC GGGGAAAATG GCACCGGCCA

541 AAGCCATCAT AACGTCTTCC CTGATGGGAA ACCTTTTCCT CACCACCCCG GATGGAGAAG

601 ATGGAATTTC ATCTACGTCT TCCACACACT TGGTCAGTAT TTCCAGAAAT TGGGACGATG

661 TTCAGTGAGA GTTTCTGTGA ACACAGCCAA TGTGACACTT GGGCCTCAAC TCATGGAAGT

721 GACTGTCTAC AGAAGACATG GACGGGCATA TGTTCCCATC GCACAAGTGA AAGATGTGTA

781 CGTGGTAACA GATCAGATTC CTGTGTTTGT GACTATGTTC CAGAAGAACG ATCGAAATTC

841 ATCCGACGAA ACCTTCCTCA AAGATCTCCC CATTATGTTT GATGTCCTGA TTCATGATCC

901 TAGCCACTTC CTCAATTATT CTACCATTAA CTACAAGTGG AGCTTCGGGG ATAATACTGG

961 CCTGTTTGTT TCCACCAATC ATACTGTGAA TCACACGTAT GTGCTCAATG GAACCTTCAG

1021 CCTTAACCTC ACTGTGAAAG CTGCAGCACC AGGACCTTGT CCGCCACCGC CACCACCACC

1081 CAGACCTTCA AAACCCACCC CTTCTTTAGG ACCTGCTGGT GACAACCCCC TGGAGCTGAG

1141 TAGGATTCCT GATGAAAACT GCCAGATTAA CAGATATGGC CACTTTCAAG CCACCATCAC

1201 AATTGTAGAG GGAATCTTAG AGGTTAACAT CATCCAGATG ACAGACGTCC TGATGCCGGT

1261 GCCATGGCCT GAAAGCTCCC TAATAGACTT TGTCGTGACC TGCCAAGGGA GCATTCCCAC

FIG. 12W

1321 GGAGGTCTGT ACCATCATTT CTGACCCCAC CTGCGAGATC ACCCAGAACA CAGTCTGCAG

1381 CCCTGTGGAT GTGGATGAGA TGTGTCTGCT GACTGTGAGA CGAACCTTCA ATGGGTCTGG

1441 GACGTACTGT GTGAACCTCA CCCTGGGGGA TGACACAAGC CTGGCTCTCA CGAGCACCCT

1501 GATTTCTGTT CCTGACAGAG ACCCAGCCTC GCCTTTAAGG ATGGCAAACA GTGCCCTGAT

1561 CTCCGTTGGC TGCTTGGCCA TATTTGTCAC TGTGATCTCC CTCTTGGTGT ACAAAAAACA

1621 CAAGGAATAC AACCCAATAG AAAATAGTCC TGGGAATGTG GTCAGAAGCA AAGGCCTGAG

1681 TGTCTTTCTC AACCGTGCAA AAGCCGTGTT CTTCCCGGGA AACCAGGAAA AGGATCCGCT

1741 ACTCAAAAAC CAAGAATTTA AAGGAGTTTC TTAAATTTCG ACCTTGTTTC TGAAGCTCAC

1801 TTTTCAGTGC CATTGATGTG AGATGTGCTG GAGTGGCTAT TAACCTTTTT TTCCTAAAGA

1861 TTATTGTTAA ATAGATATTG TGGTTTGGGG AAGTTGAATT TTTTATAGGT TAAATGTCAT

1921 TTTAGAGATG GGGAGAGGGA TTATACTGCA GGCAGCTTCA GCCATGTTGT GAAACTGATA

1981 AAAGCAACTT AGCAAGGCTT CTTTTCATTA TTTTTTATGT TTCACTTATA AAGTCTTAGG

2041 TAACTAGTAG GATAGAAACA CTGTGTCCCG AGAGTAAGGA GAGAAGCTAC TATTGATTAG

2101 AGCCTAACCC AGGTTAACTG CAAGAAGAGG CGGGATACTT TCAGCTTTCC ATGTAACTGT

2161 ATGCATAAAG CCAATGTAGT CCAGTTTCTA AGATCATGTT CCAAGCTAAC TGAATCCCAC

FIG. 12X

2221 TTCAATACAC ACTCATGAAC TCCTGATGGA ACAATAACAG GCCCAAGCCT GTGGTATGAT

2281 GTGCACACTT GCTAGACTCA GAAAAAATAC TACTCTCATA AATGGGTGGG AGTATTTTGG

2341 TGACAACCTA CTTTGCTTGG CTGAGTGAAG GAATGATATT CATATATTCA TTTATTCCAT

2401 GGACATTTAG TTAGTGCTTT TTATATACCA GGCATGATGC TGAGTGACAC TCTTGTGTAT

2461 ATTTCCAAAT TTTTGTATAG TCGCTGCACA TATTTGAAAT CATATATTAA GACTTTCCAA

2521 AGATGAGGTC CCTGGTTTTT CATGGCAACT TGATCAGTAA GGATTTCACC TCTGTTTGTA

2581 ACTAAAACCA TCTACTATAT GTTAGACATG ACATTCTTTT TCTCTCCTTC CTGAAAAATA

2641 AAGTGTGGGA AGAGACAAAA AAAAAAAAA // (SEQ ID NO. 101)

MECLYYFLGFLLLAARLPLDAAKRFHDVLGNERPSAYMREHNQLNGWSSDENDWNEKL
YPVWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCQKEDANGNIVYE
KNCRNEAGLSADPYVYNWTAWSEDSDGENGTGQSHHNVFPDGKPFPHHPGWRRWNFIY
VFHTLGQYFQKLGRCSVRVSVNTANVTLGPQLMEVTVYRRHGRAYVPIAQVKDVYVVT
DQIPVFVTMFQKNDRNSSDETFLKDLPIMFDVLIHDPSHFLNYSTINYKWSFGDNTGLFVS
TNHTVNHTYVLNGTFSLNLTVKAAAPGPCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDEN
CQINRYGHFQATITIVEGILEVNIIQMTDVLMPVPWPESSLIDFVVTCQGSIPTEVCTIISDPT
CEITQNTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPASP
LRMANSALISVGCLAIFVTVISLLVYKKHKEYNPIENSPGNVVRSKGLSVFLNRAKAVFFPG
NQEKDPLLKNQEFKGVS* (SEQ ID NO. 102)

FIG. 12Y

CTGACCAGGAACCCACTCTTCTGTGCATGTATGTGAGCTGTGCAGAAGTATGTGGCTG
GGAACTGTTGTTCTCTAAGGATTATTGTAAAATGTATATCGTGGCTTAGGGAGTGTGG
TTAAATAGCATTTTAGAGAAGAAAAAAAAAAAAAAAAAAAACTCGAGAGTACTTCTAG
AGCGGCCGCGGCGCCATCGATTTTCCACCCGGGTGGGGTACCAGGTAAGTGTACCCAA
TTCGCCTATAGTGAGT

(SEQ ID NO. 103)

AGGACAAGCCAAGGACACTCTAAGTCTTTGGCCTTCCCTCTGACCAGGAACCCACTCT
TCTGTGCATGTATGTGAGCTGTGCAGAAGTATGTGGCTGGGAACTGTTGTTCTCTAAG
GATTATTGTAAAATGTATATCGTGGCTTAGGGAGTGTGGTTAAATAGCATTTTAGAGA
AGACATGGGAAGACTTAGTGTTTCTTCCCATCTGTATTGTGGTTTTTACACTGTTCGTG
GGGTGGACACGCTGTGTCTGAAGGGGAGGTGGGGGTCACTGCTACTTAAGGTCCTAGG
TTAACTGGGGGAGATACCACAGATGCTCAGCTTTCCACATAACATGGGCATGAACCAG
CTAATCACACTGAA

(SEQ ID NO. 104)

GGATCCTTCTCCTGGTCTCCTCGGAAGAACGGGGCTTTCGCGTGACTGAGGAGAACAC
TCAGGCCCTTGCCCTTGACCGTGTTCCTGGGGCAGTTTCCTATTGGCTTGTACGCCTTG
TGTTTTTTGTACAGCAAGATGGTAACCATGGTGACAAGCACAGCCAGGCAGCCGATGG
AGATCAGGACACCATTCACTGCTCTCAGAGGGAGTCTGGGTCTTTGCCAGGGATAGAG
ATCAGGGTGCTGGTGAGGGCCAGGCTTCGATCATCTCCCAGAGTGAAATTCACACAGT
AGGTGCCAGACCCATTGAAGGCTCTTCTCACAGACAGCAGCACAGCCCATCCACAGCC
ACAGGGCTGCAGACCCGGTTCTGGGCGATCTGGCAGGTGGGGTCGGAGATGATCGTA
CAGGCTTCCATGGGGGTGGCCCCTTTGCAGGTCACAGTGAAGTCCATCAGGGAGTTGG
CAGGCTGCGGTGTGGGCATGGGGACATCTGCTATCTGCATGATGCTGACTTCCAGGATCC

(SEQ ID NO. 105)

TAGCAGATGTCCCCATGCCCACACCGCAGCCTGCCAACTCCCTGATGGACTTCACTGT
GACCTGCAAAGGGGCCACCCCCATGGAAGCCTGTACGATCATCTCCGACCCCACCTGC
CAGATCGCCCAGAACCGGGTCTGCAGCCCTGTGGCTGTGGATGGGCTGTGCTGCTGTC

FIG. 12Z

TGTGAGAAGAGCCTTCAATGGGTCTGGCACCTACTGTGTGAATTTCACTCTGGGAGAT
GATCGAAGCCTGGCCCTCACCAGCACCCTGATCTCTATCCCTGGCAAAGACCCAGACT
CCCTCTGAGAGCAGTGAAT (SEQ ID NO. 106)

GGATCCTTCTCCTGGTCTCCTCGGAAGAACGGGGCTTTCGCGTGACTGAGGAGAACAC
TCAGGCCCTTGCCCTTGACCGTGTTCCTGGGGCAGTTTCCTATTGGCTTGTACGCCTTG
TGTTTTTTGTACAGCAAGATGGTAACCATGGTGACAAGCACAGCCAGGCAGCCGATGG
AGATCAGGACACCATTCACTGCTCTCAGAGGGAGTCTGGGTCTTTGCCAGGGATAGAG
ATCAGGGTGCTGGTGAGGGCCAGGCTTCGATCATCTCCCAGAGTGAAATTCACACAGTA
(SEQ ID NO. 107)

TTTTTTTTTTTTTTTTTTAGACTGCCTTTTTAATGAGTAGAATATGTACACACACGCACC
ATACACAAAGCCCGGGCCCATTATAATTTTGTCAGGAGCTCAGGCATGCTCAGTGAGT
TGGAAGGCAGATGAAGCATG
CCTTCAGGTGGTGATTAGCTGGGTTCATGCCCATGTTATCGTGGAAAGCTGAGGCATC
TGTGGTATCTCCCCCAGTTAACCTAGGACCTTAAGTAGCAGTGACCCACCTCCCTTCAG
ACACAGCG

(SEQ ID NO. 108)

GGATCCTGGAAGTCAGCATCATGCAGATAGCAGATGTCCCCATGCCCACACCGCAGCC
TGCCAACTCCCTGATGGACTTCACTGTGACCTGCAAAGGGGCCACCCCCATGGAAGCC
TGTACGATCATCTCCGACCCCACCTGCCAGATCGCCCAGAACCGGGTCTGCAGCCCTG
TGGCTGTGGATGGGCTGTGCTGCTGTCTGTGAGAAGAGCCTTCAATGGGTCTGGCACC
TACTGTGTGAATTTCACTCTGGGAGATGATCGAAGCCT

(SEQ ID NO. 109)

FIG. 12AA

TTTTTTTTTTTTTTTTTTTTCTTCTCTAAAATGCTATTTAACCACACTCCCTAAGCCACGA
TATACATTTTACAATAATCCTTAGAGAACAACAGTTCCCAGCCACATACTTCTGCACA
GCTCACATACATGCACAGAAGAGTGGGTTCCTGGTCAGAGGGAAGGCCAAAGACTTA
GAGTGTCCTTGGCTTGTCTGGAGCAATGGATCCTTCTCCTGGTCTCCTCGGAAGAACG
GGCTTT (SEQ ID NO. 110)

AAACTGCAGTGCCCGGGCCATGCCCTCCCCCTTCGCCTTCGACTCCGCCTCCACCTTCA
ACTCCGCCCTCACCTCCGCCCTCACCTCTGCCCACATTATCAACACCTAGCCCCTCTTT
AATGCCTACTGGTTACAAATCCATGGAGCTGAGTGACATTTCCAATGAAAACTGCCGA
ATAAACAGATATGGCTACTTCAGAGCCACCATCACAATTGTAGAGGGGATCCTGGACG
CAGCATCATGCAGATAGCAGATGTCCCATGCCCACACCGCAGCCGTCCAACTCCTGAT
GGACTTCACTGTGACCTCAAGGGCACCCATGGAAGCTGTCAGA (SEQ ID NO. 111)

CCTCAACGACTCTGCCATTTCCTACAAGTGGAACTTTGGGGACAACACTGGCCTGTTT
GTCTCCAACAATCACACTTTGAATCACACTTATGTGCTCAATGGAACCTTCAACCTTAA
CCTCACCGTGCAAACTGCAGTGCCCGGGCCATGCCCTCCCCCTTCGCCTTCGACTCCGC
CTCCACCTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTG (SEQ ID NO. 112)

CCTCAACGACTCTGCCATTTCCTACAAGTGGAACTTTGGGGACAACACTGGCCTGTTT
GTCTCCAACAATCACACTTTGAATCACACTTATGTGCTCAATGGAACCTTCAACCTTAA
CCTCACCGTGCAAACTGCAGTGCCCGGGCCATGCCCTCCCCCTTCGCCTTCGACTCCGC
CTCCACCTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTGCCCACATTATCAACACCT
AGCCCCTCTTTAATGCCTACTGGTTACAAATCCATGGAGCTGAGTGACATTTCCAATG
AAAACTGCCGAATAAACAGATATGGCTACTTCAGAGCCACCATCACAATTGTAGAGG
GGATCCTGGAAGTCAGCATCATGCAGATAGCAGATGTCCCCATGCCCACACCGCAGCC
TGCCAACTCCCTGATGGACTTCACTGTGACCTGCAAAGGGGCCACCCCCATGGAAGCC
TGTACGATCATCTCCGACCCCACCTGCCAGATCGCCCAGAACCGGGTCTGCAGCCCTG

FIG. 12BB

TGGCTGGGATGGGCTGCTCCTGTCGTGGAGAAGACCTCAATGGGTCTGGTACCTACTGTGTGAATTCACTCTGGAGATGATCCAACCCT

(SEQ ID NO. 113)

GGATCCCCTCTACAATTGTGATGGTGGCTCTGAAGTAGCCATATCTGTTTATTCGGCAG
TTTTCATTGGAAATGTCACTCAGCTCCATGGATTTGTAACCAGTAGGCATTAAAGAGG
GGCTAGGTGTTGATAATGTGGGCAGAGGTGAGGGCGGAGGTGAGGGCGGAGTTGAAG
GTGGAGGCGGAGTCGAAGGCGAAGGGGGAGGGCATGGCCCGGGCACTGCAGTTTGCA
CGGTGAGGTTAAGGTTGAAGGTTCCATTGAGCACATAAGTGTGATTCAAAGTGTGATT
GTTGGAGACAAACAGGCCAGTGTTGTCCCAAAGTTCCACTTGTAGGAATGGCAGAGTC
GTTGAGG

(SEQ ID NO. 114)

CCTCAACGACTCTGCCATTTCCTACAAGTGGAACTTTGGGGACAACACTGGCCTGTTT
GTCTCCAACAATCACACTTTGAATCACACTTATGTGCTCAATGGAACCTTCAACCTTAA
CCTCACCGTGCAAACTGCAGTGCCCGGGCCATGCCCTCCCCCTTCGCCTTCGACTCCGC
CTCCACCTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTGCCCACATTATCAACACCT
AGCCCCTCTTTAATGCCTACTGGTTACAAATCCATGGAGCTGAGTGACATTTCCAATG
AAAACTGCCGAATAAACAGATATGGCTACTTCAGAGCCACCATCACAATTGTAGAGG
GGATCCTGGAAGTCAGCATCATGCAGATAGCAGATGTCCCCATGCCCACACCGCAGCC
TGCCAACTCCCTGATGGACTTCACTGTGACCTGCAAAGGGGCCACCCCCATGGAAGCC
TGTACGA

(SEQ ID NO. 115)

GAAGGTGGAGGCGGAGTCGAAGGCGAAGGGGGAGGGCATGGCCCGGGCACTGCAGTT
TGCACGGTGAGGTTAAGGTTGAAGGTTCCATTGAGCACATAAGTGTGATTCAAAGTGT
GATTGTTGGAGACAAACAGGCCAGTGTTGTCCCCAAAGTTCCACTTGTAGGAAATGGC
AGAGTCGTTGAGGAAGTGGCTGGGATCATGAATGAGGACATCGAAGACGA

(SEQ ID NO. 116)

FIG. 12CC

GAATTCGCACGAGGGGAGTCAGAGTCAAGCCCTGACTGGTTGCAGGCGCTCGGAGTC
AGCATGGAAAGTCTCTGCGGGGTCCTGGGATTTCTGCTGCTGGCTGCAGGACTGCCTC
TCCAGGCTGCCAAGCGATTTCGTGATGTGCTGGGCCATGAACAGTATCCCGATCACAT
GAGAGAGCACAACCAATTACGTGGCTGGTCTTCGGATGAAAATGAATGGGTTCCAATA
TCACTTTTGTGGTGAA

(SEQ ID NO. 117)

GAATTCGGCACGAGGAAGGAGGCCGTGTGCAGGCAGTCCTGACCAGTGACTCACCGG
CTCTGGTGGGTTCCAATATCACTTTTGTGGTGAACCTGGTGTTCCCCAGATGCCAGAAG
GAAGATGCTAATGGCAATATCGTCTATGAGAAGAACTGCAGGAATGATTTGGGACTG
ACATCTGACCTGCATGTCTACAACTGGACTGCAGGGGCAGATGATGGTGACTGGGAAG
ATGGCACCT

(SEQ ID NO. 118)

GAAGGTGGAGGCGGAGTCGAAGGCGAAGGGGGAGGGCATGGCCCGGGCACTGCAGTT
TGCACGGTGAGGTTAAGGTTGAAGGTTCCATTGAGCACATAAGTGTGATTCAAAGTGT
GATTGTTGGAGACAAACAGGCCAGTGTTGTCCCCAAAGTTCCACTTGTAGGAAATGGC
AGAGTCGTTGAGGAAGTGGCTGGGATCATGAATGAGGACATCGAAGACGATGGGGAG
GTCTCTGAGGAAGATCTCATCAGACAAGTT

(SEQ ID NO. 119)

GAATTCGGCACGAGGTCAAGCCCTGACTGGTTGCAGGCGCTCGGAGTCAGCATGGAA
AGTCTCTGCGGGGTCCTGGGATTTCTGCTGCTGGCTGCAGGACTGCCTCTCCAGGCTGC
CAAGCGATTTCGTGATGTGCTGGGCCATGAACAGTATCCCGATCACATGAGAGAGCAC
AACCAATTACGTGGCTGGTCTTCGGATGAAAATGAATGGATGAACACCTTGTATCCA

(SEQ ID NO. 120)

FIG. 12DD

AAGGGGGAGGGCATGGCCCGGGCACTGCAGTTTGCACGGTGAGGTTAAGGTTGAAGG
TTCCATTGAGCACATAAGTGTGATTCAAAGTGTGATTGTTGGAGACAAACAGGCCAGT
GTTGTCCCCAAAGTTCCACTTGTAGGAAATGGCAGAGTCGTTGAGGAAGTGGCTGGGA
TCATGAATGAGGACATCGAAGACGATGGGGAGGTCTCTGAGGAAGATCTCATCAGAC
AAGTTCCTGTCATTCTTCTGGGACATGGTCACGAATACAGGGATCTGATCTGTTAT

(SEQ ID NO. 121)

GAATTCGGCACGAGCCGACACTGTGACTCCTGGTGGATGGGACTGGGGAGTCAGAGT
CAAGCCCTGACTGGTTGCAGGCGCTCGGAGTCAGCATGGAAAGTCTCTGCGGGGTCCT
GGGATTTCTGCTGCTGGCTGCAGGACTGCCTCTCCAGGCTGCCAAGCGATTTCGTGAT
GTGCTGGGCCATGAACAGTATCCCGATCACATGAGAGAGCACAACCAATTA

(SEQ ID NO. 122)

AAGGTGAAAGATGTGTATGTGATAACAGATCAGATCCCTGTATTCGTGACCATGTCCC
AGAAGAATGACAGGAACTTGTCTGATGAGATCTTCCTCAGAGACCTCCCCATCGTCTT
CGATGTCCTCATTCATGATCCCAGCCACTTCCTCAACGACTCTGCCATTTCCTACAAGT
GGAACTTTGGGGACAACACTGGCCTGTTTGTCTCCAACAATCACACTTTGAATCACAC
TTATGTGCTCAATGGAACCTTCAACCTTAACCTCACCGTGCAAACTGCAGTGCCCGGG
CCATGCCCTCCCCCTTCGCCTTCGACTCCGCCTCCACCTTCGTA (SEQ ID NO. 123)

TACCATCGGAGAAAGAAGACCAAGCAAGGCTCAGGCAGCCACCGCCTGCTTCGCACT
GAGCCTCCTGACTCAGACTCAGAGTCCAGCACAGACGAAGAGGAATTTGGAGAATTG
GAAATCGCTCTCGTTTTGTCAAGGGAGACTATCCCGATGCTGCAAGATCTGCTGTCCCT
CTGGCCTTTGTCATCCTCGCGCCTGCGTTGTGGCCTCTGTGGGCTTGGTGTGGAGCAAA
TGGCTCTCAAGGAGGACTGAGTCTCAAGGAAATT (SEQ ID NO. 124)

FIG. 12EE

AGCTAAGGTCAGGAGGTGTCTGAAGAATTGGCTGATGCATGGCAGGGATGTTGTTGAC
CTGCTTTTAGAACAATACTTCCATTTAATTATAGCATATCTTATGTGTGTATTAAAGCA
GAGCCGATCTGGTGGGGCTCATTAAGTAAATGTACTTACTGCAAAAGGTTCAACTGGT
GACCCCAGTTTTCCCCAGAAGCAATATGATAGGACAGAGGCGACTCCTGCAAGTTGTC
TCAGACTTCACACATACATTGTGACATTCTCTGAGCATGTGCACTGTACATGATATGAC
ACTATCAA

(SEQ ID NO. 125)

AGCTAAGGTCCACTACCTTGTGAAGATGTATAAACACCTGAAATGTAGAAGCGATCCG
TATGTCAAGATCGAGGGGAAGGACGCTGACGACTGGCTGTGTGTGGACTTTGGGAGTA
TGGTGATCCATTTGATGCTTCCAGAAACCAGAGAAACCTATGAATTAGAGAAACTATG
GACTCTACGTTCTTTTGATGACCTTAGCTAAGCCGAATCAGCACACTGGCGGCGTTACT
AGTGGATCGAGCTCGTACAGCTGATGCATAGCTTGAGTATCTATAGGTTACTAATAGC
TGGCTATCATGTCAAGCGTTC

(SEQ ID NO. 126)

GCTGAGCTGCAGAGAGTAGCACATCCTTGCTAATTCAATAACTACCAGTTTTTATTGGT
GAAACATGAATCCAGATGGTATGGTTGCTCTCCTGGACTACCGTGAAGATGGTGTGAC
TCCATTCATGATTTTCTTTAAGGATGGCTTAGAGATGGAGAAATGTTAACAAATTGGA
TCTATCACCTGTCACCATAATTGGCTGCTGCTTACCATCCATACAACACCAGGACTTAG
GACAAATGGGACTGATGTCATCTTGAGCTTTTATTTTGACCTTAGCT

(SEQ ID NO. 127)

AGCTAAGGTCAGAGCCAATAGTATCATGAGAACTGAAGAAGTAATAAAGCAACTTCT
CCAGAAATTTAAGATTGAGAATAGCCCTCGGGATTTCGCTCTTTACATTATTTTTGGGA
CAGGAGAGCAGAGAAAGCTAAAGAAGACCGATGTCCACTGCTGCAGAGGTTACTACA
AGGACCATCCAAAAGCAATGCTCGGATCTCTCATGGATAAAGATGCAGAAGAATCAC
GAGAGATGTGGCTCGTACATTATTTCACTTTCTTCTGATCATACTCAAGATAGATGAGA
GAGAAT

(SEQ ID NO. 128)

FIG. 12FF

TTGACTTCTGAGTCTAACACAGACACTGCAAGGGTTAATTTTCCAAGAGGTGGTTGTT
GTTGACGATAAATTCATTAAGAATTTTTAAAAATTTAGTTAGATTTACCAAAGTCACTG
GAGACAAATTCAGAAGGCATATATACCTGCCAGTTTTGTGGACTACATTAATAGGGAG
GCTTTTATGTTTGATGTAATTCTTACAGTTCTAAGAATTAAGTTCCATTGCATGAGACC
TTAGCT (SEQ ID NO. 129)

AAGGTGAATCCCCGACGGCTCTGGGCCCGAGGAGAAGCGTCGCCGTGGCAAATTGGC
ACTGCAGGAGAAGCCCTCCACAGGTACTTGGAAAAACTGGTCTCTGAGGCCAAGGCC
AGCTCCGAGACATTCAGGACTTCTGGATCAGCCTCCAGGGACACTGTGCAGTGAGAAG
ATGGCCATGAGTCCTGCCAGTGAG (SEQ ID NO. 130)

AATTTTTTTTTTCGACGGCCCAACGGGGGCTTGGTGGATGGAAATATGGTTTTGTGAGT
TATTGCACTACCTGGAATATCTATGCCTCTTATTTGCGTGTACTGTTGCTGCTGATCGT
TTGGTGCTGTGTGAGTGAACCTATGGCTTAGAAAAACGACTTTGTCTTAAACTGAGTG
GGTGTTCAGGG (SEQ ID NO. 131)

CACCTGATTTAAAGGAAAAGCATTCTGACGTAAGAAGCTGAAAGGCGGCCCTTGCGTG
CTTTGAACTTTCTTATACAGCACAGTCATCTGAAGCTTCCTGTGTGACCAAGACAAGA
ACGCGTGCACAAGACTGAGAAACAGCAAGAAACAACCCGGCATTCTACTTTCTCAAC
ACTATCATACTTTAAACCTTTCAC (SEQ ID NO. 132)

FIG. 12GG

CTAGCTTACGCTAGTCCCCCATGCATAAAGACTGATCGCTTTTCCTTAGAAAGGTGAG
AGGGTTAGGACAAGGCCGTGTGGTAACAACACCCGCAGCTCGAAAAACCAATGGCTT
GTTAACGTGTCAGTGAGGCACTGTACGGACGTCCATAGTCCACATCTTCAAATTCCCG
CAGAAGGCTTCCTATTCTTAAACTCTA

(SEQ ID NO. 133)

CTACATTTCTGTATCCATTCCTCTGTTGAAGGCTCTGGTTCTTTCCAGCTTCTGGCTATT
ATAAATAAGGCTGCTATAAACACAGTGGAGGCATGTGTCCTTGTTATATTTTGGAGCA
TCTTTTGGGTATATGCCCAGAAGTGCTATAGCTGGTTCCTCAGGTAGTACTATGTCGAA
TTTTCTGAGGAACTGCCAGACTGATTTCCAGAGTGGTTGTACCAGCTTGCAATCCCACC
AGCAATAGAGGAGTGTTCCTCTTTCTCTATATTCTTGCCAACATCTGCTGTCACCTGAG
TGTTT

(SEQ ID NO. 134)

TGGTAAAGGGGGAATGATGTCGAGGCCATCCTGGGCTGTAGAGCCAGGCCCTGGCTTG
GGGAGTGGGCATTGTTAACTTGTTGCTGACTTTGTGTTGACCCCTGCATCAGCAACTAT
TTCCTTAAATCCAGGATACAACTTGTTAAGTGTGACAGCTTTCCTTTACACACCATTTT
TGTGGGTGTATATATATATTTGACTTGGGGAGAATTATTTTTTACAAAAATACAAAAT
AGCTTTTAA

(SEQ ID NO. 135)

AGCTAAGGTCCGGACTCTATGGCATGACCCCAAAAACATTGGCTGGAAAGATTACACT
GCCTACAGGTGGCACCTGATTCACAGGCCTAAGACAGGCTACATGAGAGTCTTAGTGC
ATGAAGGAAAGCAAGTCATGGCTGACTCAGGACCAATTTATGACCAAACCTACGCTG
GTGGACGGCTGGGCTGTTTGTCTTCTCCAAGAGATGGTCTATTCTCGGACCTCAAGTAT
GAGTGCAGAGATGCTAGAGAGCAGGCTCAGTCTCAGCA

(SEQ ID NO. 136)

FIG. 12HH

TGACCTACGTGTAGTTGGTGTGCTTGTTGTCGAAGATGAGGGCCTCCTGGATGAGCTG
GTGCTGCTGCTCCAGCAGGTCCAGGCTGGGCTTGTAGTCCACGAGTCTGCGCTCGTAC
TGCTTCAGGTGGCTCAGCTGGTCTTCCAGAGTCCCGTTCATCTCAATGGAGATGCGCCC
GATCTCCTCCATCTTAGTCTGGATCCACGGCCCCACCATATTGGCTTGGCTGGCGAACT
GTCGGCGAAGGCTGCATTGGATTGCT

(SEQ ID NO. 137)

AATTTTTTTTTTCGACGGCCCAACGGGGGCTTGGTGGATGGAAATATGGTTTTGTGAGT
TATTGCACTACCTGGAATATCTATGCCTCTTATTTGCGTGTACTGTTGCTGCTGATCGT
TTGGTGCTGTGTGAGTGAACCTATGGCTTAGAAAAACGACTTTGTCTTAAACTGAGTG
GGTGTTCAGGG

(SEQ ID NO. 138)

CACCTGATTTAAAGGAAAAGCATTCTGACGTAAGAAGCTGAAAGGCGGCCCTTGCGTG
CTTTGAACTTTCTTATACAGCACAGTCATCTGAAGCTTCCTGTGTGACCAAGACAAGA
ACGCGTGCACAAGACTGAGAAACAGCAAGAAACAACCCGGCATTCTACTTTCTCAAC
ACTATCATACTTTAAACCTTTCAC

(SEQ ID NO. 139)

CTAGCTTACGCTAGTCCCCCATGCATAAAGACTGATCGCTTTTCCTTAGAAAGGTGAG
AGGGTTAGGACAAGGCCGTGTGGTAACAACACCCGCAGCTCGAAAAACCAATGGCTT
GTTAACGTGTCAGTGAGGCACTGTACGGACGTCCATAGTCCACATCTTCAAATTCCCG
CAGAAGGCTTCCTATTCTTAAACTCTA

(SEQ ID NO. 140)

CTACATTTCTGTATCCATTCCTCTGTTGAAGGCTCTGGTTCTTTCCAGCTTCTGGCTATT
ATAAATAAGGCTGCTATAAACACAGTGGAGGCATGTGTCCTTGTTATATTTTGGAGCA
TCTTTTGGGTATATGCCCAGAAGTGCTATAGCTGGTTCCTCAGGTAGTACTATGTCGAA

FIG. 12II

TTTTCTGAGGAACTGCCAGACTGATTTCCAGAGTGGTTGTACCAGCTTGCAATCCCACC
AGCAATAGAGGAGTGTTCCTCTTTCTCTATATTCTTGCCAACATCTGCTGTCACCTGAG
TGTTT

(SEQ ID NO. 141)

TGGTAAAGGGGGAATGATGTCGAGGCCATCCTGGGCTGTAGAGCCAGGCCCTGGCTTG
GGGAGTGGGCATTGTTAACTTGTTGCTGACTTTGTGTTGACCCCTGCATCAGCAACTAT
TTCCTTAAATCCAGGATACAACTTGTTAAGTGTGACAGCTTTCCTTTACACACCATTTT
TGTGGGTGTATATATATATTTGACTTGGGGAGAATTATTTTTTACAAAAATACAAAAT
AGCTTTTAA

(SEQ ID NO. 142)

AGCTAAGGTCCGGACTCTATGGCATGACCCCAAAAACATTGGCTGGAAAGATTACACT
GCCTACAGGTGGCACCTGATTCACAGGCCTAAGACAGGCTACATGAGAGTCTTAGTGC
ATGAAGGAAAGCAAGTCATGGCTGACTCAGGACCAATTTATGACCAAACCTACGCTG
GTGGACGGCTGGGCTGTTTGTCTTCTCCAAGAGATGGTCTATTCTCGGACCTCAAGTAT
GAGTGCAGAGATGCTAGAGAGCAGGCTCAGTCTCAGCA

(SEQ ID NO. 143)

TGACCTACGTGTAGTTGGTGTGCTTGTTGTCGAAGATGAGGGCCTCCTGGATGAGCTG
GTGCTGCTGCTCCAGCAGGTCCAGGCTGGGCTTGTAGTCCACGAGTCTGCGCTCGTAC
TGCTTCAGGTGGCTCAGCTGGTCTTCCAGAGTCCCGTTCATCTCAATGGAGATGCGCCC
GATCTCCTCCATCTTAGTCTGGATCCACGGCCCCACCATATTGGCTTGGCTGGCGAACT
GTCGGCGAAGGCTGCATTGGATTGCT

(SEQ ID NO. 144)

TGACCATCGATAAGTTTAATAACTACAGACTTTTCCCAAGACTACAAAAGCTTCTTGA
AAGTGACTACTTTAGATATTACAAGGTGAACTTGAAGAAGCCTTGTCCTTTCTGGAAT

FIG. 12JJ

GACATCAACCAGTGTGGAAGAAGAGACTGTGCCGTCAAACCCTGCCATTCTGATGAAG
TTCCTGATGGAATTAAGTCTGCCGAGCTACAAGTATTCTG
AGGAAGCCCAACCGCATTGAAGAATGTGAGCAAGCTGAGCG (SEQ ID NO. 145)

AACTCTGTGAACCGTGCCTTTCTCTGTGGAGGTGGAGGTGTCGGTTGAAGACAAGCGA
GGTCCTCCAAGGGGCTGTGTCTTATGTTGCCATCTCCCCTTGTAGCTTGGCTGCCCACC
CTCCAGACTGTGCGCCATGGCTCCAAGGCTGTGACCCGCCACTGGAGTCATGCACTTC
CAGCGGCAGAAGCTGATGCTATAACTGAGTATATTCCTCCAAACCTGCCATCAACCCG
AGA (SEQ ID NO. 146)

ACTTCTCCAGAGAATTTAAGATTGAGAATAGCCCTCGGGATTTCGCTCTTTACATTATT
TTTGGGACAGGAGAGCAGAGAAAGCTAAAGAAGACCGATGTCCCACTGCTGCAGAGG
TTACTACAAGGACCATCCAAAAGCAATGCTCGGATCTTCCTCATGGATAAAGATGCAG
AAGAAATCAGCAGAGATGTGGCTCCGTACATTAATTTCACTTTTCTTTCTTGGATCCAT
CCTTCAAGATTAGATGAAGAAGAGAAATGGAGATTGAGAGAATATGCAATCATACCGA
(SEQ ID NO. 147)

AGGGTTACTTCAGGCTAAGGCAATAGAAATCCATTTTAAGATGGTGTGCTAAAGGCTT
GATGGATGTTCATCGTCTGTCTAAAGGAGAATGAAGTCATCAACAGGATGTCAGGGGA
AAGTGAGATCATCGCAGAAAGTATCAACTTAGCACAAACACACAGGCATAGCTCCTG
CAAGAGGTGAATGCTGTCCCCAAATACCTGAGGAACTATCCCTTTGGGCAAGAAAATA
GACAAGTCCATGAAGTCTGGGTGA
(SEQ ID NO. 148)

GACCAGGTACACTTGAGCAAAGCACCCAGTATTTAATTCCTTACAGAAAGGAGAGGA
AAGGTCTGCAGTTGGACTGATGGTATGCTAACACCGCAAATGACTGTCATTTGATCTC

FIG. 12KK

AGAAGTTCAGGATTGATTGCTATGTTTTAGCTCTAATTGTGAGAAACAGTAGTCATTTT
AGTCTTAAATTTTGCCCTCAGGAAATTCAGGGAGACTGAGCCTTCCTTCCCCCACCTTC
GTAAAGCCGAATTCCAGCACACGGCGGCCGTTACTAGTGGATCCGAGCTCG

(SEQ ID NO. 149)

TACAAGGTGGGATGGCAGGAACTGAAGGCTTCTGTAAATCCAGTTTTGGCTCTCTCTC
TGGTCTTTCTTTCTCTTCTGTTCTGTTTGGAAGGGTTTCTGGTCTTTCAGGAGGTATTTT
TTTAATTTCATGTTTTCTCTCTGTGGTACCTGCCCCTTGTTTGACGACAGGAGCTGATG
GAGGTGGCGGTTTCTTGGGTCTATTCCCTTCCTTGTCAAAGTCCGATGGAAGTAACTTC
ACGAAGTTGTCAGGAAACACGCCTCGTCTGCCATTGAGTTCTCCTTCCCACCAGCCTA
CGCGATGCAGTCTTATTGATGAGAGTCACTATATCTCCTTA (SEQ ID NO. 150)

TCACCCATGACTTCTATGGACTTGTCTATTTTCTTGCCCAAAGGGATAGTTCCTCAGGT
ATTTGGGGACAGCATTCACCTCTTGCAGGAGCTATGCCTGTGTGTTTGTGCTAAGTTGA
TACTTTCTGCGATGATCTCACTTTCCCCTGACATCCTGTTGATGACTTCATTCTCCTTTA
GACAGACGATGAACATCCATCAGGCCTTTATGCACACCATCTTAAAATGGATTTCTAT
TGCCTTAGCCTGAAGTCC

(SEQ ID NO. 151)

CCCATAGAGATAGGTTTGCTCCAGAACCTGCAGCATTTGCACATCACAGGGAACAAGG
TGGACATTCTGCCAAAACAGTTGTTTAAGTGCGTGAAGTTGAGGACTTTGAACCTGGG
GCAGAACTGTATCGCCTCCCTGCCTGAGAAAATCAGTCAGCTCACCCAGCTCACTCAG
CTGGAGCTGAAGGGCAACTGCCTAGACCGCCTGCCAGCCCAGCTGGCAGTGTCGATGC
TCAAGAAGA

(SEQ ID NO. 152)

CAATAATCCAGGTAAAATAGAGTAAAATAGTCTGCTAGCAGCAAGTTCCTACCATACT
TTCAACAACACTCACGAGATACGGAATGATTACAGCATTAAGAATATTTCAGAAATGA
CAGGTAGGTGTGGTGGACAGGTGGCTCACATTCAAGACTCAAGTCTACTTAAAAAAGA

FIG. 12LL

AAATCTCACTAGCACTAGATTCTAGCTCCTTTGTTTCCCCCTTTCTTTTGGTTTCAAAG
GCGTTTCTACAACCCATAAGAGG

(SEQ ID NO. 153)

GCCAAGCTATTATGACACTATAGATACTCAACGTATCGATCAACGTTGGTACCGAGCT
CGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGGCTTGGATTGGTCAGAGCA
GTGTGCAATATGATCCAACTAAGTCTCCTCCCTTGGCCCCTCCCCAAAATGTTTGCAGT
GTTATTTTTGTGGGTTTTTTTTTAACACCCTGACACCTGTTGTGGACATTGTCAACCTTT
GTAAGAAAACCCAAATAAAAATTGAAAAATAAAATAAAAAGAAACCCATGAACATTC
GCACCACTTGTGGCTTCTGACTATCTTCCACAGAGGGAAGTTTAAAACCCAAACTTCC
AAAGGTTTGAACTACCTCAAGACACTTTCGCAGTGGAGTCGTAGACCAATCCCA

(SEQ ID NO. 154)

TAAATAAATTAAAAAACTATTAAACCTAAAAACGTCCACCAAACCCTAAAACCATTAA
ACAACCAACAAACCCACTAACAATTAAACCTAAACCTCCATAAATAGGTGAAGGCTTT
AATGCTAACCCAAGACAACCAACCAAAAATAATGAACTTAAAACAAAAATA

(SEQ ID NO. 155)

GGTAAAGGGGACCTGGAGAACGCCTTCCTGAACCTGGTCCAGTGCATCCAGAACAAG
CCCCTGTACTTCGCTGACCGGCTGTACGACTCCATGAAGGGCAAGGGGACTCGAGACA
AGGTCTGATTAGAATCATGGTCTCTCGCAGTGAAGTGGACATGCTGAAAATCAGATCT
GAATTCAAGAGGAATATGGCAAGTCCTGTACTACTACAT (SEQ ID NO. 156)

AGAGCAGCAGGCCAGCTGTACTTGGTTTGGCAAGAAAAAGAAGCAGTACAAAGATAA
ATATTTGGCAAAGCACAACGCAGTGTTTGATCAATTAGATCTTGTCACATATGAAGAA
GTAGTCAAACTGCCAGCATTCAAAAGGAAAACATTAGTCTTATTAGGTGCACATGGTG
TTGGAAGAAGACACATAAAAAATACCCTCATCACAAAGCAC (SEQ ID NO. 157)

FIG. 12MM

TCGGTCATAGTAGTAAGGGAAATCTCCCAGGTAAGATGAATACTGCGGTAGGACGAA
CAATCCTCCAGGATGTTTGTTCCATATTAAACTGTTACGTGATATGTGCTTGAATATTC
TGTCCTGAATAATCTCTAGTGTAGTTAATACAATCTTCTCAACTGAAGAAAAATAAGC
CTCCCACAAGAACTGTGTCTGCTGTCTAAGTGCTAGGATTTTATCCTGATGAATAGACC
TGATTGTAGAAGGAATCTGTAATAGCAATCTCTCATCGCCTATGACCGAAAGCCGAAT
TCTGCAGATATCCATCACACTGGCCGGCCGCTCGAGCATCGATCTAGAGGG

(SEQ ID NO. 158)

CTGCTTGATGACAAAGGGTGTAGTCTTCATCTTTTCCTGGATTATTTTGGAAGTGACAG
GTGGAAATTCCATCGTCACGTTTATGTGGTCTGTAAAGCCAACGATCTCAAATTCTGG
CGGCTCAAGAGGAGCGTTTGCAGGCACGATGTAGTCTGAGCAGCGGCACACGGTCAA
GTCCCCTCTGTGCACTATGACGATGGCGACGACGTAGCTCTCCATGCCCTCCAACCAC
TTATCTGTCACGTCACATGATGACTTCGTGGTATCTGAACAGTTCTTAACCTTCGTCAG
ATTTTCGTCTTT

(SEQ ID NO. 159)

AAATCGTTGCTTCAGAAAGACTCAATAACACTTACTTGTGCCTGGCTGTGCTGACAGT
ACATTCTGTGTCATTTTCCTTCATGGGCGGAACAGTCCACAGAGCTCACCAACAAGTA
CTCCAAAACTGAGCAAGAGTTTAAGCTTCGAGATGCAACCAGATGAGCTTCTAGAAAA
GCCCATGTCTCCCATGCAGTACGCACGGTCTGGACTAGGGACAGCAGAGATGAATGGC
AAACTCATAGCTGCAGGTGGTTATAACAGAGAGGAATGTCTTCGAACAGTTGAATGCT
ATGATCCACATACAGATCACTGGTCCTTCCTTGCTCCCATGAGAACATCAAGCAG

(SEQ ID NO. 160)

CTTTCCGAAGAGCACACCCTCCTCTCAATGAGCTTGTGAGGTCTCTTTCTTCTCTTCCT
TCCAACGTGGTGCTAGCTCCAGGCGAGCGACGTGAGAGTGCCACCTGAGACAGACAC
CTTGGTCTCAGTTAGAAGGAAGATGCAGGTCTAAGAGGAATCCCCGCAGGTCTGTCTG
AGCTGTGATCAAGAATATTCCGCAATGTGCCTTTTCTGAGATCGTGTTAGCTCCAAAG

FIG. 12NN

CTTTTTCCTATCGCAGAGTGTTCAGTTTGTGTTTGTTTGTTTTTGTTTTGTTTTGTTTTTC
CCTTGGCGGATTTCCCGTGTGT (SEQ ID NO. 161)

CCTATTGAACGGTCTTGCAATGACGAGCATTCAGATGCTTAAGGAAAGCATTGCTGCT
ACAAATATTTCTATTTTTAGAAAGGGTTTTTATGGACCAATGCCCCAGTTGTCAGTCAA
AGCCGTTGGTGTTTTCATTGTTTAAAATGTCACCTATAAAACGGGCATTATTTATGTTT
TTTTTCCCTTTGTTCATATTCTTTTGCATTCCTGATTATTGTATGTATCGTGTAAAGGAA
GTCTGTA (SEQ ID NO. 162)

CCTATTGAACGGTCTTGCAATGACGAGCATTCAGATGCTTAAGGAAAGCATTGCTGCT
ACAAATATTTCTATTTTTAGAAAGGGTTTTTATGGACCAATGCCCCAGTTGTCAGTCAA
AGCCGTTGGTGTTTTCATTGTTTAAAATGTCACCTATAAAACGGGCATTATTTATGTTT
TTTTTCCCTTTGTTCATATTCTTTTGCATTCCTGATTATTGTATGTATCGTGTAAAGGAA
GTCTGTA (SEQ ID NO. 163)

CCTGGGTCCGTCCTCCAACCCCTCACGCCCAAACCCTCCGACTTTCACTTCTTGAAGTG
ATCGGAAAGGGCAGTTTTGGAAAGGTTCTTCTGGCTAGGCACAAGGCAGAAGAAGTA
TTCTATGCAGTCAAAGTTTTACAGAAGAAGCCATCCTGAAGAAGAAAGGAAGGAAGC
ATATTATGTCAGAGCGGAATGTTCTGTTGAAGAATGTGAAGCACCCTTTCCTGGTGGG
CCTTCACTTCTCATTCCAGACCGCTGACAAGCTCT (SEQ ID NO. 164)

GATGCTGAACACAAAAAGAAAGAAGAAAAGGAAGAGGAGGAGCAAGAGAAGCTGAA
GGGAGGGAGCCTTGGCGAAAATCAGATCAAAGATGAGAAGATTAAAAAGGACAAAG
AGCCCAAAGAAGAGTCAAGAGCTTCTTGGATAGAAAGAAAGGATTTACAGAGTGAGG
CGCAGAATGGAGATTCATGACCCACAAACTTAAAC (SEQ ID NO. 165)

FIG. 12OO

AAAGCCAATTGGTAGAGAAATTGAAGACACAAATGCTGGATCAGGAAGAGCTTCTGG
CATCAACCAGAAGGGATCAAGATAATATGCAAGCTGAACTGAATCGCCTCCAAGCAG
AAAATGATGCTTCTAAAGAAGAGTAAAGAGTTTTACAGGCCTTAGAGGACTGCTGTTA
ATTATGATCAGAGTTCAGGAGTTAAGAC

(SEQ ID NO. 166)

CTGCTTGATGTCCTGTGTAGCGAATGTCACAGCGTACAACATTGTTAGTGTAGTCTGAT
TCAGGCACCAGGTAGCTGGGGTTTACACTGACCTTTAGAATGTAGTTTCCAGGTTGTA
CATCTGTAATATCAATCCACTGGCAGTCTATGTCTGCCGCATAGGTGTCATAACATCCA
GGACTCAATCCCTGTGTGTGTGCAGTGCACGCAAAGGCCCTGTGGTACCCATAGTCAC
AGGACGTGTCCTCCAGACAGAAGCTTGCTTTGTGGCCTTCAGCCACTCTCCTCTGTGTG
TTGGCATCAACGAGAAGCCGAATTCTCGAGATATCCATCACACT (SEQ ID NO. 167)

CTGCTTGATGTCCTGTGTAGCGAATGTCACAGCGTACAACATTGTTAGTGTAGTCTGAT
TCAGGCACCAGGTAGCTGGGGTTTACACTGACCTTTAGAATGTAGTTTCCAGGTTGTA
CATCTGTAATATCAATCCACTGGCAGTCTATGTCTGCCGCATAGGTGTCATAACATCCA
GGACTCAATCCCTGTGTGTGTGCAGTGCACGCAAAGGCCCTGTGGTACCCATAGTCAC
AGGACGTGTCCTCCAGACAGAAGCTTGCTTTGTGGCCTTCAGCCACTCTCCTCTGTGTG
TTGGCATCAACGAGAAGCCGAATTCTCGAGATATCCATCACACT (SEQ ID NO. 168)

GATCTGACACTACAGCATGAGCGTTAGATTTCATAAAATTATTTTTCTTCTAAATGCTG
GAAACTCTAAGGGTTTATTCAGAAAAAAAACTGGCCAATTTTCAAATGGCTTAGAAGC
AGGGTTAATTAAGTATTGAATGAGCCACTGTGATATCCTGATGACACCCAGTCACAAT
GACAGTTTTGAAGCATACAACCAAAACAATTGAGATCTCAAAACTATTTTACATCACT
TATGGTAATGTTATGTAAAAATGAAAATGCTTTCTGTGGAAGTTACATTCTTTACCAGG
TCTTTAACATAAATTAACACGACGTCGAGTAAGCCTTTGTTCGGAAGACAAACTAGTT
TGTGAGTTCAGTCAGATCCCAGCT

(SEQ ID NO. 169)

FIG. 12PP

AGTTGCCAGGACCACCACCATAGTTGCCAGGTTCATCATAAACAAATCCAACATCAAT
CTTAAATTCCCCCATCAGACAATCTGCCCTCAAAGAATGGGAATTATAAACCCGGATA
CTGATGATCTCATCCATGAGCTCAGAGGGTGTGATGTGCACATTGTAGAAAAATAACT
CGTCAAAAAACGGATTGTTCCCTCTCTTGATTCTCGTGCGATGCGTCTGACCACAGATG
TGAACTTTCACCACGGGCCTTATGTTGTTGCCGCATAACTGACGGCCCTCGATCACTCT
GACACGGATCTGGAAATCTGTGGCTTGTTGGACAGCATCCTT (SEQ ID NO. 170)

AAGCCGTGTCCCAAAGAATGGATAGAGACGCGATCAGATGCGACAGTGCTGTGGAGA
AAGCCCAGGAACCTGCACAATTGCCCTGGTCCAATGGCTCGTGGATCAGGTTGGGCCA
CTTCTCTGAAGCTTCAAAGGCAGTGGGTAGCACTTCCCCTTGGCCCAGCACCGTATAA
ATCTCATTCATATTCATGACAGTGGAGGATGGGCGGATTGTGCCCAGGCGGTACGGAA
TGCCCTCATCCAGGGTCATGCCCCAGAAGGCACTGTGGTTCCCAGCCTGCCACCCGTA
GTTGCCTCGGTTGATGGCTTTAATCATGTCTGGTCACTAGACACGGCTTAAGCGAATCT
CGAGATATCCATCACACTGGCGGCGTCGAGAT (SEQ ID NO. 171)

AAGCCGTGTCTGATGATGGAGGTAGTGGTGGGGGAGGAGGGACTGAGGGTCCTGAGG
TGGTGGCCCCTGGAACTGATCCCACATAGTTACCCACTGCTAGTTCTGACCCCGTGGA
CAACGTGCCAGAGGCCATGACTGGCAGTATGGCAATGTCCCCATCCCCTTTCTTCTTA
ATTTTAATGGTCCCTTGTTTCTCCAGTTCGTGAATCTTTTTTTCCAGGGTAGACTGTCTT
TGAATGGCTTCTTCCTTTTCTTTGACCATTTTTCTTAACGTGTGAACTTGGGTATTTGCA
TCTTTGTAGATTTCCGGACAACATCAGTTCCTTATTCCTCTGCATAAGTTGCTTTCAGTT

(SEQ ID NO. 172)

CGAGTCAGACACATGAAAGCAAAACGCGGGCAGATAAAACGATCGCCTTACCTTCTA
GCAAAAATCTGAAGCTTGTGTCAGAAACAAAGACTCAGAAAGGTTTGTTTTCAGATGA
AGAAGACTCTGAGGATTTGTTTTCTTCTCAAAGTTCAAGTAAGCCAAAAAGTGCATCA

FIG. 12QQ

CTTTCATCCAGCCAGCCCCCAACATCAGTCTCCCTTTTTGGTGATGAAGATGAAGAGG
ACAGTCTTTTTGGGAGTGCAGCAGCTAAGAAGCAGACTTCATCTCTACAACCTCAGAG
TCAAGAGAAAGCAAAGCCTTCCGAGCAGCCCTCAAAGAAGACATCTGCCTTGTTGTTC
AGA

(SEQ ID NO. 173)

CGAGTCAGACTTAATTTAAAAACGAAACAAAACAAAAATAACATAGTTTAGAAATCA
AGGAGAAAGGACAGATAGTCTAAGAAAAAAGACAACACAAAAGAGGGGCAGGGCGG
CCAGCTTGCATCAGGGATCTTGGCTGGAGACCTGCTTTGAATAGGTTTCTTGCAGGTAT
TTCTTAAATGCTGTGGGGTTTTTCCAGAGTTCCGCAGCGTGTGTGTTCAAAGGGCTATC
GATGTTGGGTTCTCCTAGCAGGCTCTGGATAGAGAGCAAGATAGTCCTGACATCATAT
AGTGCAGACCACTTATCCTTGAGGATGTCCGGCAGATGTTGCCTGGGTGTCACGTTGG
GGTGGTAGCAGGGTGTGAGGAACTTCACTG

(SEQ ID NO. 174)

CGAGTCAGACACTCCTGGCTCCTGGATTCTTTAGATGCCTCCATCAGACTGGGTACTTT
AGATGCCTCCATCAGACTACTTCGTCATTGTATTTCTCAGTTCGCTCAGGGCAAGCGGC
AGTCTCTGGGCTGCTGTGGCAGGTGCCACCACTGCATTTAAAAGTTAAAATTTCTTCA
AATATTCCCATCAAGGCCTTGTAGCCTCTGAGATTGGTTTACTATTTGCCCAGTTATTT
AAAGCTCTCTGCATTCCTTCCTGATTTAATATTGCTATGGCCAGGACAATGTGTAGAAG
TAAAAAGGATATCATATTTACAGGTGTAACGC

(SEQ ID NO. 175)

FIG. 12RR

DD-PCR PRIMER AND PCR SIZE (nt)	cDNA FROM CELL LINE	MOUSE HOMOLOGY (%nt)	HUMAN HOMOLOGY (%nt)	NORTHERN (P-MT) (SCREEN 1)	NORTHERN-CLONED DNA (P-MT) (SCREEN 2)
P17-6 cl10 (1100)	151-1 LM1	MUSCLE NICOTINIC ACETYLCHOLINE RECEPTOR APLHA (54.3%)		NO	151-1LM1 UP, 151-1LMA DOWN
P19-6 cl2 (500)	151-2 PA		LYMPHOCYTE IgE RECEPTOR (52.6%)	NO	151-2LMA DOWN,DOWN
P21-6 cl3 (450)	151-2 PA	HISTON H2b (94.2%)		151-1LM1 DOWN,DOWN	151-1LM1 DOWN,DOWN
P21-9 cl6 (500)	151-1 PB	RATTUS NORVEGICUS THIOL- SPECIFIC ANTIOXIDANT mRNA(94.4.%)		151-1LM1 DOWN,DOWN 151-2LMA UP,UP	151-1LM1 DOWN,DOWN 151-2LMA UP,UP
P21-17 cl9 (1000)	148-1 LMD	MUS MUSCULUS PUTATIVE PROTEIN TYROSIN PHOSPHATASE mRNA(98.3%)		148-1LMD UP,UP 151-1LM1 UP,UP	148-1LMD UP,UP 151-1LM1 UP,UP
P22-5 cl3 (600)	148-1 LMD	RAT DIHYDROPYRIDINE-SENSITIVE L-TYPE CALCIUM CHANNEL ALPHA-2 SUBUNIT GENE (92.5%)		148-1LMD UP,UP	148-1LMD UP,UP

FIG. 13A-I

P22-6 cl4 (600)	148-1 LMD	SAME AS P22-5 Cl3		148-1LMD UP 151-1LM1 UP	148-1LMD UP,UP
P22-9 cl3 (800)	148-1 LMD	RAT KIDNEY ZN-PEPTIDASE AMINOPEPTIDASE N mRNA (90.5%)		148-1LMD UP,UP,UP	148-1LMD UP,UP,UP
P24-6 cl3 (550)	151-1 PB		UBIQUITIN CARRIER PROTEIN (E2-EPF) mRNA (53.3%)	151-1LM1 DOWN 151-2LMA UP 151-2LMB UP	151-2LMA UP
P24-10 cl3 (1400)	151-1 LM1	RATTUS NORVEGICUS CALPAIN II 80 kDa SUBUNIT mRNA (93%)		151-1LM1 UP,UP	151-1LM1 UP,UP
P25-1 cl3 (400)	148-1 PA	M. MUSCULUS KERATINOCYTE GROWTH FACTOR Fgf-7 (99.4%)		148-1LMD DOWN 151-1LM1 DOWN,DOWN 151-2LMB UP,UP 151-2LMMA UP	148-1LMD DOWN 151-1LM1 DOWN 151-2LMB UP 151-2LMMA UP
P25-9 cl8 (1300)	151-1 PB	M. MUSCULUS mRNA FOR INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN-3(98.1%)		148-1LMD UP 151-1LM1 DOWN,DOWN,DOWN 151-2LMA UP,UP,UP	148-1LMD UP 151-1LM1 DOWN,DOWN,DOWN 151-2LMA UP,UP,UP
P2-27 (cl18-3)	148-1 PA	RATTUS NORVEGLUS GLYPICAN mRNA (93.4%)			148-1LMD DOWN P53(+)12 DOWN

FIG. 13A-2

CLONE #	cDNA FROM CELL LINES	DD PRIMER	PCR SIZE (nt)	MOUSE HOMOLOGY	HUMAN HOMOLOGY	NORTHERN BIG BIOT #	REGULATION TYPE	SEQUENCING PRIMER	SEQUENCING LENGTH
CI 3#1 CI 4#1 (SAME FRAG & ORIENTATION	151-2 LMB	P3		TYROSINE KINASE? VIP2	CAVEOLIN (70%)	N123 148-1 UP 151-1 UP 151-2 UP	UP	-40	241 156
CI 5A#4	148-1 PA	P2		THROMBO-SPONDIN 100%	THROMBO-SPONDIN	N124 148-1 DOWN 151-1 DOWN 151-2 UP	DOWN	-40	233
CI 25#3	151-2 LMA	P5			53BP2 P53-BINDING PROTEIN (53.3%)	148-1 DOWN 151-1 DOWN 151-2 UP	DOWN		
CI 29#3 CI 28#1 (SAME FRAG; DIFFERENT ORIENTATION	148-1 LMD	P5	335 332		TGF-BETA 2 (53.0.%) Kvi-1 nmls(53.0%)	N119 148-1 UP 151-1 UP 151-2 UP	UP	T7	335 332
CI 54A#2	141-1 PA	P8		MUSCULUS RECEPTOR TYROSIN KINASE CYCLIN G	PROTO-ONCOGENE TYROSINE PROTEIN KINASE GENE	N126 148-1 DOWN (WEAK) 151-1 DOWN (WEAK) 151-2 UP (WEAK)	DOWN	Sp6	220

FIG. 13B-1

FIG. 13B-2

CI 63#4	151-2 LMA	P10			Y316 GENE (53.8%) 1AC GENE (53.8%) Rb SUSCEPTI-BILITY GENE (50%)	N127	UP	Sp6	340
CI 74#2	151-2 LMA	P11#3		86.8% SERUM & GLUCOCORTICOID REGULATED KINASE (sgk)		N120 148-1 UP 151-1 DOWN 151-2 UP		Sp6	320
CI 75#1	151-2 LMA	P11#10		87% MATCH sgk			UP	Sp6	250
CI 78B#4 MATCH THE SAME GENE BUT DIFF. FRAG.	148-1 LMD	P12		92.2% MATCH sgk	PROTEIN KINASE C-L (57%)			Sp6	270

DD-PCR PRIMER AND PCR SIZE (nt)	MOUSE HOMOLOGY(%nt)	HUMAN HOMOLOGY (%nt)	TGF-BETA STIMULATORY RESPONSE (12 hr.)	NORTHERN (P-MT)	CELL LINE
P11-2 cl5 (310)	LYSYL OXIDASE (100%)		⇧⇧⇧	⇩⇩	N132: 148-1 LMD, 151-1 LM1 DOWN, 151-2 LMB, 151-2 LMC UP
P20-23 cl9 (850)	ACTIN BINDING PROTEIN(100%)		⇧⇧	⇧⇧	N142: 148-1 LMD, 151-2 LMA,LMB,MMA UP, 151-1 LM1 UNCHANGED
C129-3 (P5) (335)		NMB(79.8%)	⇩⇩	⇧⇧	N119: 148-1 LMD 151-1 LM1, 151-2 LMA,LMB,LMC,MMA UP
P17-3 cl8 (1000)	UBIQUITIN ACTIVATING ENZYME E1(100%)		⇧	⇩⇩	N142: 151-2 LMA DOWN
P20-3 (400)		ALPHA ACTININ 3 mRNA (77.5%)	⇧⇧		
P18-12 cl3 (1000)	RAT mRNA FOR P34 PROTEIN (89.6%)		⇧		
P25-7 cl3 (1000)	M.MUSCULUS mRNA FOR P19-PROTEIN TYROSINE PHOSPHATASE (100%)		⇧	⇧⇧	148-1IMD UP
P19-1 cl3 (310)		POLYMORPHIC LOCI IN Xq28 (30%)	⇧		

FIG. 13C

DD-PCR PRIMER AND PCR SIZE (nt)	MOUSE (RODENT) HOMOLOGY (%nt)	HUMAN HOMOLOGY (%nt)	SCREEN 1 P53 STIMULATORY RESPONSE (12h. OR 24h.)	SCREEN 2 CLONED DNA
P1-8 cl10 (1000)		DYSTROPHIN GENE (50.4%)	P53(+)24 DOWN,DOWN	P53(+)24 DOWN,DOWN
P1-9 cl10 (500)	M.MUSCULUS mRNA FOR CYCLIN G (96.5%)		P53(+)12 UP,UP P53(+)24 UP,UP,UP	P53(+)12 UP,UP,UP P53(+)24 UP,UP,UP
P7-4 cl1 (600)	RATTUS NORVEGLOUS SGK mRNA (51.3%), RAT LUNG DERIVED LO1 C-ros-1 PROTO-ONCOGENE mRNA (48.4%)	NITRIC OXIDE SYNTHASE (47.1%)	148-1LMD DOWN P53(+)12 UP,UP P53(+)24 UP,UP,UP	P53(+)12 UP P53(+)24 UP
P9-17 cl9 (500)	RAT mRNA FOR CYCLIN D1 (79.1%)		P53(+)24 UP	P53(+)24 UP
P9-20 cl3 (850)		H. SAPIENS LDLC mRNA (51.8%)	P53(+)12 DOWN P53(+)24 DOWN,DOWN	P53(+)24 DOWN
P11-23 cl2 (800)	SYRIAN HAMSTER GENE FOR CYTOCHROME P-4 (52.5%), RAT CARBOHYDRATE BINDING RECEPTOR GENE (50.6%)		P53(+)24 UP,UP	P53(+)24 UP
P15-9 cl1 (600)	MOUSE (CLONE BALB11N) mRNA (47.2%)	PTGS2 GENE FOR PROSTAGLANDIN ENDOPEROXIDE SYNTHASE-2 (46.6%)	P53(+)24 DOWN	P53(+)24 DOWN,DOWN
P15-14 cl5 (500)			P53(+)12 UP P53(+)24 UP	P53(+)24 UP
P18-23 cl10 (500)			148-1LMD DOWN P53(+)12 DOWN P53(+)24 DOWN	148-1LMD DOWN P53(+)12 DOWN P53(+)24 DOWN

FIG. 13D

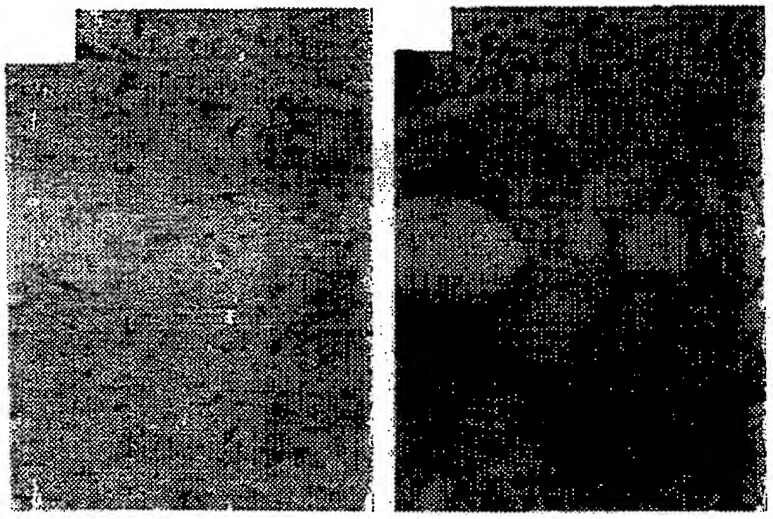

FIG. 14A FIG. 14B